Notes from the EC'13 Program Chairs

R. PRESTON MCAFEE
Google
and
ÉVA TARDOS
Cornell University

This note describes our experience as 2013 ACM Electronic Commerce (EC) program chairs and summarizes a survey we ran after the program was complete.

Track System

EC is the leading scientific conference on advances in theory, systems, and applications at the interface of economics and computer science, including applications to electronic commerce. EC'13 continued the 2012 track process involving three non-exclusive focus areas or tracks. The three tracks are:

—Theory and Foundations (TF)

—Artificial Intelligence and Applied Game Theory (AI)

—Experimental, Empirical, and Applications (EA)

We expanded the AI track to include applied game theory, on the principle that applied game theory is usually more like AI than CS theory. The track system represents a compromise. The expansion of the AI track helped more explicitly align the track with an area at the interface of economics and computation, and have the SPC dedicated to this track consist of a mix of computer scientists and economists, analogous to the mix present in the other tracks.

Authors were asked to align their submission with one or two of the tracks. The Program Committees, both senior and regular, were also associated with tracks. The papers submitted to one track were reviewed by PC and SPC members associated with that track, while a paper submitted to two tracks was handled by PC and SPC members in the union of the two tracks, with at least one PC and one SPC from each of the designated tracks.

Allowing authors to align their submission with one or two of the tracks seems to work well for the conference from our own experience, in the opinions expressed at last year's business meeting, and in the opinions of about 90% of those who responded to our survey this year. The main criticism of the complementary 10% was that the tracks overlapped, rendering them confusing. There was certainly some confusion between TF and AI that may be alleviated by SIGecom discussion. In contrast, there was little confusion about experimental and empirical work. The ability to submit to multiple tracks mitigates this confusion. Moreover, the identification of SPC with tracks reduces the ambiguity of submission, based on the SPC's

Authors' addresses: preston@mcafee.cc, eva@cs.cornell.edu

own research. We think the result of the track system is a more diverse program, specifically more papers by AI researchers, than would be created with a single track system. A mix of computer science and economics PC and SPC members in each track results in the desired cross-fertilization across computing and economics that is the hallmark of EC since its inception.

Program Committee and Assignment

Our first task as Program chair was to put together the Senior Program Committee (SPC), and the Program Committee (PC) for each track. Based on last year's submissions, we anticipated that more of the submissions will be for the theory track than either of the two other tracks. In spite of their greater numbers, TF SPCs handled considerably more papers each. In contrast, we aimed to have the PC burden more evenly split. Many of the PC members were suggested by the SPC.

Table 1	SPC	PC	Submissions	Accepted
TF	15	73	167	52
AI	8	33	54	16
EA	8	27	50	16

The call for papers attracted 223 submissions from authors in academia and industry all around the world, 2 of which were eventually retracted. Out of the 221 non-retracted submissions, 50 papers choose dual tracks.

We selected a diverse and capable Program Committee and senior Program Committee who are eminently capable of evaluating any paper in the field. We then optimized the matching of skills and the nature of papers as best we can. Because of the size of the matching problem we used an algorithm described below. Creating the best matching took us several days, but it insured that relevant expertise was available for all 221 papers. We believe that all the papers were carefully evaluated by peers with relevant skills.

Once the papers were submitted, we asked the SPC and PC to indicate the papers they could review, would like to review, or could not review due to a conflict. Both groups also provided keywords about their skills and interests. PC members were limited to blind selection of papers (based on title and abstract only) for confidentiality and asked to mark author conflicts separately.

Based on paper preferences, conflicts and keywords, we used integer programming to generate an initial assignment of papers to PC and SPC members. The main ingredients of this optimization were the constraints of the track system explained above and limits on the number of papers assigned to a PC or SPC, with a heavy weight on paper preferences and a lighter weight on keyword matches.

Each PC member was asked to review of a maximum of 8 papers. Similarly SPC papers also had hard upper bounds on the number of assigned papers depending on the track (higher in TF than in AI and EA).

To make sure each paper gets enough expert reviews, we aimed to have at least one of the assigned PC members and at least one of the assigned SPC very well aligned to each paper. We reviewed the initial assignment, tweaked it, and re-optimized many times to reach the final assignment. We increased the weight on paper preferences so that most reviewers received papers they requested, and most

papers were reviewed by reviewers that requested the paper. The additional tweak involved finding papers without adequate expertise (e.g. because of rare or missing keywords) and forcing an assignment to insure the expertise was available.

Furthermore, we aimed to have each PC and SPC member have at least half of their load come from papers that they specially asked for to make the PC/SPC work both enjoyable and to make the ratings fair, and this was accomplished in most cases.

Review Process

To help the make the ratings more uniform, we announced an interpretation of the ratings associated with journals in economics and computer science. PC members had an extremely short three weeks to review the assigned paper. We owe a special thanks to the diligence of the PC, as almost all reviews arrived by the deadline. The review period was followed by a two day author feedback period, where authors saw the reviews, and were given the opportunity to comment on the reviews and respond to issues raised by them. Authors were limited to 500 words.

Opinion in the community about the author feedback option is decidedly mixed. Many feel that author feedback is useful chance for authors to correct misconceptions by the reviewers, and clarify issues raised. Others feel that author feedback is a waste of time and energy. Author feedback certainly worked well for some papers where the PC and SPC suspected that the paper had a mistake, or where the PC and SPC had questions for the author. In cases where the Program Committee has questions for the author, the PC chairs can also serve as a conduit to obtain answers.

Once we had the reviews and the author feedback, the PC and SPC engaged in two weeks of discussion. The discussion period is the most interesting and valuable part of the EC Program Committee's job. Many of these discussions were quite lively and involved both computer science and economics perspectives, as befits our community. A few key points to comment on:

Most papers had extensive discussions. Usually this discussion is summarized by an additional review (meta-review) provided by one of the assigned SPC members, but unfortunately not all papers received such meta-reviews, which is a social waste because useful information is not communicated the authors.

The review process identified a number of errors in manuscripts, including erroneous proofs. In all cases we insisted that the last submission before the deadline be evaluated or the paper withdrawn. To replace a paper after the deadline is to let the deadline slip; it isn't fair to let some replace and not others. The author feedback system allowed authors to comment on small mistakes, as well as misunderstandings.

It would be desirable to reduce the number of errors in submissions. One way to achieve this is for authors to seek feedback from their colleagues before submission, and avoid the "just in time" production process favored by many computer scientists.

The discussions, involvement of the SPC, and author feedback had a major effect on the final decision. Prior to the start of the discussion period, few SPC members were involved in reviewing. Their role was to lead the discussion and guide the

decision made on the paper. A good number of reviewers adjusted their ratings during the discussion phase, but not all discussions are reflected in such changed ratings. The bigger effect of the discussion was how it shaped the final decision.

The quality of the reviews prompted the most comments in our survey. Overall, most respondents felt that the reviews were reasonable. Only 8% of the respondents were very unsatisfied with the reviews, while 33% were very satisfied. Providing high quality and useful feedback to authors is a hard task both due to the compressed time line of the EC review system, and the broad background of reviewers and authors and the broad range of areas of the conference.

One form of complaint expressed by some in the survey is that reviews, ratings, and decisions are too random. While there is certainly some randomness involved in evaluating papers, the PC members have substantial agreement on overall quality. To assess the agreement empirically, we considered the absolute variation from the mean in the "overall" rating. The three PC members submitted their scores without knowing each other's scores, so these are independently submitted. The mean was 5.47, and average deviation (absolute value) from this mean was 1.54. Thus, the average score was 1.54 away from the global mean of 5.47. In contrast, the average deviation from each paper's mean was only 0.92. So ratings of the same paper were significantly closer to each other than ratings of different papers. After the discussion, and including the SPC rating, these numbers are 1.47 and 0.81, respectively.[1]

Table 2	Initial	Final
Average Absolute Deviation	1.54	1.47
Average Within Paper Deviation	0.92	0.81

An important source of unpredictability in evaluating papers is what the reviewers find interesting. This difference was especially acute when reviewers have very different backgrounds (some are economists, others are computer scientists) because the two disciplines have different perspectives on what is interesting. This mix of economics and computer science in the discussions, as well as the mixed audience of the conference, is designed to result in the desired cross-fertilization across the disciplines.

Program Size and Structure

We decided to keep roughly the same number of parallel and single sessions as last year, accepting 72 papers in the program (same as last year), and the roughly 30% acceptance rate (on each of the three tracks). We also added two keynote speakers. We asked the SPC for suggestions and after some discussion and votes by the SPC, we asked Jon Kleinberg and Al Roth to give keynote talks at EC'13 and both accepted.

The community has strong opinions on what is the right number of papers to accept for EC. Using double sessions for part of the time, and a very compressed

[1]While absolute deviations make the deviation interpretation easier, they make it harder to interpret how much of the variation is explained by paper quality. We find that mean paper quality explains 57% of the total variation in scores, rising to 58% after the discussion. Random chance would produce 33% and 25%, respectively.

program last year's EC'12 increased the accepted papers from 48 at EC'11 to 72 at EC'12. We were strongly encouraged by the community (by discussions at EC'12 business meeting) as well as the SIGecom leadership to keep this higher number of papers at EC'13, which we did. Our survey confirmed the same range of opinions. Out of the 149 responses, the overwhelming majority (104 responses) liked the current mix of single and double session, and the remaining 45 responses were almost evenly split between wanting all single or all parallel sessions. Overall, we think that the 30% acceptance rate and the mix of single and parallel sessions worked well for EC, though many commented that the number of accepted papers should not be increased too fast. The current state appears to be a good compromise between two competing forces:

—Higher acceptance rate helps in building the community, and allows more people to give talks, allows a greater mix of talks, and broader range of topics.

—Lower acceptance rate increases the prestige of the conference and helps increase the quality relative to other venues.

As in the past years, authors of accepted papers can ask that only a one page abstract of the paper appear in the proceedings, along with a URL pointing to the full paper. This option is made available to authors to accommodate the publishing traditions of different fields, where a conference publication precludes an overlapping journal article. Authors were not asked indicate their plan to use this option during the same process.

The ACM EC best paper award rules are set by SIGecom, and are detailed at `http://www.sigecom.org/awardp.html`. One aspect of that was not discussed in the rules is the role of the one page papers. The SIGecom executive committee decided that only full length paper qualify for such awards. The conference can award one or two Best Paper Awards and one or two Best Student Paper Awards to the accepted papers. The Best Paper Award is made irrespective of whether or not a paper is a student paper - a Best Paper that is a student paper is also awarded Best Student Paper. This year, as a student paper won Best Paper, there is only one award.

The award selection process develops in two stages. The program chairs are asked to nominate a small subset of highly rated papers (5 papers this year), some of which are student papers, and are asked to form an Awards Committee from members of the senior Program Committee or the wider research community, who do not have conflicts of interest with the nominated papers. This year's award committee consisted of Susan Athey (Stanford), Vincent Conitzer (Duke), David Easley (Cornell), and Anna Karlin (University of Washington). The winner of the best paper award will be announced at EC in Philadelphia.

Last Word

We believe that the upcoming EC'13 program looks great, and want to thank all the people who contributed to make this happen. We are indebted to Kevin Leyton-Brown and Panos Ipeirotis, the 2012 chairs, for their generous assistance and advice, to Thomas Preuss of Confmaster, and Pooya Jalaly Khalilabadi for their help with running the review system. We want to thank all the authors who

submitted papers, the Program Committee and the Senior Program Committee for their contribution to this process, as well as SIGecom chair David Parkes for his continued helpful advice throughout the process.