Strategyproofness in Peer Review

https://j-b-z.github.io/

Jeremy Zhang (jbz)

November 2018

1 Project Description

To do my research in strategyproof proof peer review, I will be working with Assistant Professor Nihar Shah (https://www.cs.cmu.edu/ nihars/) and PhD students Yichong Xu (http://xycking.wixsite.com/yichongxu) and Han Zhao (https://www.cs.cmu.edu/ hzhao1/). Faculty member Nihar Shah is an assistant professor in the ML and CS departments. Yichong and Han are fourth year PhD students.

1.1 Motivation

To understand the motivation behind research in peer-review, it is critical to first recognize the significance of conferences themselves. For the most part, conferences represent a medium for researchers to communicate their ideas and findings in a particular field. However, conferences also have a critical and often understated impact in a researcher's career. Papers have shown that having even a single paper accepted to a conference can severely affect a young researcher's chances to be accepted into prestigious research groups and universities.

Since conferences themselves have such a crucial position in the current academic environment, we naturally expect the most out of the peer-review systems which select the papers accepted into conferences. However, other studies have exposed many weaknesses within current peer-review systems ranging from biases and tendencies within the reviewers, to inherent disadvantages for papers whose research topic is niche or novel, to strategies to game and rig the results of peer-review. By researching strategyproofness in peer-review, researchers can understand more about this particular aspect of peer-review, and hopefully create a system that cannot be gamed.

1.2 Methods and Previous Work

Peer review itself is typically separated into two approaches: a cardinal approach (reviewers give a score to papers reviewed) and an ordinal approach (reviewers will rank the papers reviewed). In both approaches, the results are aggregated and then made into a complete ranking. My research will focus on the ordinal approach.

Within the ordinal approach, strategyproof peer review is analyzed with two terms: $strategyproofness$ and $efficiency$. $Strategyproofness$ itself is not unique to the peer review setting. In fact, it can be studied in many social situations. However, $strategyproofness$ in peer review captures the intuition that a reviewer cannot improve the ranking of their own paper in the complete ranking by altering their personal rankings. Rigorously, $strategyproofness$ is the property that for every pair of paper and reviewer such that the reviewer is an author for the paper, then for any assignment (aggregation of reviewer rankings) which differs by only the rankings provided by the specific reviewer, the ranking of the specific paper should be fixed.

$Efficiency$ is a heuristic used to measure how accurate and useful the complete rankings are, relative to the individual reviewer's ranks. The definition of $efficiency$ typically varies from paper to paper. Even so, $efficiency$ is typically defined from the lens of statistics or social choice. Typically, the statistical definition of $efficiency$ would measure the similarity of reviewer's ranking to the complete rankings. On the other

hand, a social choice definition of $efficiency$ would be a rigid property that may or may not be satisfied by the complete rankings with respect to reviewer's rankings.

Shah recently published a paper on strategyproof peer review systems. In this paper, he defines two notions of efficiency under a social choice lens called group unanimity and pairwise unanimity. Group unanimity is satisfied if for every partition of the papers such that every reviewer reviewing papers from both sets consistently ranks papers from the first set over the second, then the complete rankings ranks papers from the first set over the second. Pairwise unanimity is satisfied if for every pair of papers such that every reviewer that reviews both papers consistently ranks the first paper over the second, then the complete rankings ranks the first paper over the second. For these two notions of efficiency, we note that pairwise unanimity implies group unanimity. Thus, pairwise unanimity is stronger than group unanimity.

With these two notions of efficiency, Shah constructs a partitioning algorithm and shows that this algorithm guarantees that a disconnected graph can be partitioned such that the resulting assignments satisfy are strategyproofness and and group unanimity. Then, Shah shows several negative results, including impossibility to satisfy pairwise unanimity. Furthermore, Shah hypothesizes that strategyproofness might not be achieved if the graph is connected. Shah concludes his research with several avenues to explore as a follow up. These avenues are the areas I will specifically be exploring in my research.

1.3 Research Proposal

My research revolves around extending upon Shah's most recent paper. The questions I seek to answer are:

Are strategyproofness and group unanimity mutually exclusive when the authorship graph is connected?

Can a weaker notion of strategyproofness allow for more efficiency?

Can the addition of reviewers with no conflicts lead to significant improvements to efficiency?

Can the elimiation of specific authors lead to better partioning algorithms?

How efficient is the partitioning algorithm given a statistical definition of efficiency?

2 Project Goals

Inherently, my research consists of several theoretical open questions. By its nature, progress may be difficult to judge and measure. Even so, listed below are expectations for my project.

75%: Even if I encounter many roadblocks in my research, I expect to make headway in several of the proposed research questions. Specifically, I should be able to rigorously analyze at least two of the proposed questions and provide substantial and meaningful write-ups for my progress and analysis in these questions.

100% If all goes well, I expect to finish analysis on several of the proposed research questions. Specifically, I should be able to sufficiently analyze and provide conclusions to at least two of the proposed questions. The write-ups for these questions should be significant enough to publish.

125% If research proceeds faster than expected, I expect to achieve the results as in the 100% goal, as well as perhaps explore some avenues that I personally find interesting in regards to strategyproof peer review.

3 Milestones

3.1 1st Technical Milestone

By the end of the semester, I hope to gain rudimentary intuition for each of the proposed research questions, and then choose three of the questions to specifically pursue in depth. Furthermore, I should have all the literature search papers read.

3.2 Bi-weekly Milestones

Again, due to the theory-based nature of research, the timelines and milestones may be wildly subject to change.

Week 2: Read research papers specifically relevant to the chosen research questions and discuss with Nihar, Yichong, and Han.
Week 4: Discuss ideas for exploring the chosen research questions/discuss potential for research reconsiderations (assess and address speed of research).
Week 6: Discuss other research papers specifically relevant to the chosen research questions.
Week 8: Discuss additional areas of interest.
Week 10: Discuss additional areas of interest.
Week 12: Wrap up research and summarizing of findings.
Week 14: Review of write-up.

4 Literature Search

This list is still bare but will expand as I continue to look into the field:

- Xu, Y., Zhao, H., Shi, X., Shah, N. B. (2018). On Strategyproof Conference Peer Review. arXiv preprint arXiv:1806.06266.
- Stelmakh, I., Shah, N. B., Singh, A. (2018). PeerReview4All: Fair and Accurate Reviewer Assignment in Peer Review. arXiv preprint arXiv:1806.06237.
- Hajek, B., Oh, S., and Xu, J. (2014). Minimax-optimal inference from partial rankings. In Advances in Neural Information Processing Systems, pages 1475–1483.
- Fischer, F. and Klimm, M. (2015). Optimal impartial selection. SIAM Journal on Computing, 44(5):1263– 1285.

5 Resources Needed

At the moment, I am not aware of any hardware/software requirements for my research, since it seems to be theory-oriented.

Paper resources such as other technical papers in peer-review seem to be available online and accessible for free.