Game Theoretical Modeling and Studies of Peer-Reviewing Methods

Matthew John PETERSON and Marius C. SILAGHI
Florida Tech, USA, Email: {petersom,msilaghi}@fit.edu
Makoto YOKOO
Kyushu University, Japan, Email: yokoo@is.kyushu-u.ac.jp

Abstract

We model a family of peer-reviewing processes as game-theoretic problems. This helps to understand elements of existing peer-reviewing procedures, and to predict the impact of new mechanisms. Real world conference peer-reviewing processes are overly complex and here we define and analyze simple versions. Our versions provides a contribution towards formalizing and understanding these complex real world problems.

The players of this game are *the researchers* that participate as authors and reviewers. *A funding agency* tries to maximize the social value by providing rewards to researchers based on their publications. Equilibria of truthful reviewing in described mechanisms is analyzed experimentally.

1 Introduction

Social mechanisms sometimes assume that politicians, electoral commissions, and peer-reviewers are altruistic. Research in the game theory community has intensively analyzed and improved decision making in auctions [10] and voting processes [2, 6] Just as in the case of politics, money from funding agencies can transform reviewers into self-interested parties with quantifiable utilities. We now introspectively turn our scrutiny unto our own performance in the complex decision process involved in the peer-reviewing for evaluation of scientific results.

Conference peer-reviewing mechanisms are becoming rich in features, in an attempt to provide more accurate evaluations, to encourage valuable research, and to raise the quality of the conferences. While we are not aware of earlier work modeling and analyzing peer-reviewing as a game-theoretic problem, the conference peer-reviewing process is undergoing essential transformations. The need to analyze it with game theoretic approaches is stressed in [3]. Features at fashion include:

- blind reviewing: assuring the secrecy of the reviewers' names, and (with double-blind reviewing) even secrecy of the authors' names,
- enabling authors to give answers to reviewers' comments,
- enabling reviewers to bid for papers,
- enabling authors to rate reviewers, and
- enabling authors to blacklist reviewers with potential conflicts of interest.

We model the quality of an article (or text) by its utility to the society, which we call *worth*. The worth quantifies how much the community gains or loses from the publication of that article. The worth is evaluated by expert reviewers, and defines the *types* of these reviewers. The type t of a reviewer is a function $t : \Gamma \rightarrow \mathbf{R}$ specifying the worth of each article, where Γ is the set of articles and $\mathbf{R}$ is the set of real numbers. Since all reviewers are considered equally expert, they have the same type.

In order to model the utilities of each author and reviewer, we assume that they expect rewards from *funding agencies*. We assume that the funding agencies intend to maximize a *social value*, assumed to be defined based on the total quality of the published papers. Given a set of texts Γ appearing in a community of type t, the social choice function (which defines the desired outcome) is

$$f^{\Gamma}(t) = \text{“publish} \quad A = \{\gamma | \gamma \in \Gamma, t(\gamma) > 0\}\text{”}$$

maximizing the total utility $\sum_{\gamma \in A}(t(\gamma))$.

A researcher can get *bad reputation* by publishing erroneous articles (i.e., articles with a negative worth), or when any expert can publicly verify that the researcher performed an incorrect review.

We assume that, prior to seeing reviewing comments, a researcher does not know the worth of his own articles (explaining erroneous articles).

In one setting we assume that since the funding agency cannot directly measure the utility (worth) of a text, it distributes funding based on a *citation influence* computed based on the number of citations from papers presented in conferences. Therefore, cheating strategies by researchers can be motivated by a desire to increase their citation influence (funding): for example, we assume that a reviewer may gain by slowing the publication of articles that *supersede* the reviewer's work. A paper is *superseded* by a newer paper when the new paper presents a better solution, thereby reducing the influence of the older paper. In another setting, agencies offer funding based on the count of published papers.

In the next section we introduce background about game theory. After introducing in Section 3 the formal definitions of the main involved concepts, the next section discusses some of the implications of blind reviewing. Further, a commonly used double blind reviewing mechanism is formalized, as well as a non-blind reviewing mechanism suggested in [3]. Section 4 evaluates what happens if funding agencies reward scientists based on the number of publications in the conference rather than based on counting citations. In the end, an experimental comparison of the described mechanisms is presented, together with conclusions.

2 Background

In one version of the Prisoner's Dilemma (PD) [5, 4], two prisoners are asked to testify against each other in court. If they collaborate in not testifying, each gets 1 year of prison. If both testify against each other (defect), then each gets 5 years of prison. If only one prisoner defects, then he will go free and the other prisoner gets 10 years of prison. This is the one shot version of the game. If we assume that the game repeats many times (i.e., several times, after being freed, the same players will be arrested again and asked to replay the game), then we have an *iterated* game (or *repeated* game) [1]. A *payoff matrix* is a matrix used to show the utilities (penalties) of the players actions for the one-shot game. For the described penalties and players i and j, the matrix is:

	j cooperates	j defects
i cooperates	-1 , -1	-10 , 0
i defects	0 , -10	-5 , -5

In each cell there is a pair of payoffs, one for each participant: first payoff for the first participant (i), and the second for the other participant (j). By *rationality* of a player one understands the assumption that the player is interested only in his own well-being and does not care about the fate of the other player. The rational action for each prisoner (independently of what the other prisoner chooses), is to defect.

However, for the iterated version of the game, different considerations apply. A well studied strategy for repeated PD is the *forgiving Tit-for-Tat* [8]. Namely, a prisoner prefers to collaborate but if his colleague defects, then the first prisoner will punish him by defecting with a high probability in the next iteration of the game. The strategy is forgiving in the sense that a player tries to avoid an infinite cycle of Tit-for-Tat punishments by sometimes choosing collaboration (with a small probability) even after the other player defects. Forgiving Tit-for-Tat is famous for performing extremely well in average (assuming there will be an infinite number of iterations) although it is quite simple.

A rational player is a player that maximizes his expected utility. Therefore, a rational player plays repeated PD according to the best performing strategy that he knows, such as forgiving Tit-for-Tat. However, maximal common utility is obtained if both prisoners always collaborate.

A set of strategies in a game are in Nash equilibrium [7] if no player can do better by changing his strategy while the strategies of the other players remain fixed.

3 Main Concepts

Let us now specify formal definitions of the main concepts used in defining Peer-Reviewing Games (PR Games) [9].

Definition 1 (Paper Worth) *The* worth *of a paper is the utility brought to society (tax-payers) by publishing that paper.*

The issue of determining the worth of a paper is similar to the issue of determining the utility of an item in auctions, and we assume that it can be determined by any careful expert reviewer. The type of the reviewer specifies the utility evaluated for each paper. We assume that all reviewers are experts having the same type, t.

Definition 2 (Social Value) *The social value given a set of publication venues* Ψ *(posters, regular papers, etc.) consists of the weighted sum of the worth of the published papers, with weights* $w_\psi, \psi \in \Psi$*:*

$$\sum_{\psi \in \Psi} \left(\sum_{\gamma \text{ published in } \psi} (w_\psi * t(\gamma)) \right)$$

Definition 3 (Superseded Paper) *A paper* P_1 *is superseded by paper* P_2 *if authors cite* P_2 *instead of* P_1*.*

Definition 4 (Citations Influence) *The* citations influence *(CI) of an author at a given moment is a metric of the influence of his publications, and is given by the weighted sum of the worth of his un-superseded publications of each of the three venues:* $\Psi = \{$*regular, poster, technical report*$\}$*. The corresponding weights are* w_o*,* w_p *and* w_t*.*

$$CI = \sum_{\psi \in \Psi} \left(\sum_{\text{un-superseded gamma in } \psi} (w_\psi * t(\gamma)) \right)$$

Note that a publicly known erroneous text brings to its author's CI a penalty equal with the absolute value of the worth of that text (weighted with the factor defined for that type of publication), due to *bad reputation.*

We assume that conferences take place yearly, updating researchers' CIs.

Definition 5 (Funding Agency) *The* funding agency *is the entity that yearly pays each researcher an amount computed using some function based on the existing publications.*

While the funding agency cannot access the worth of a paper directly, in some scenarios we assume that it can evaluate the same CIs in a different way (e.g., by some kind of counting actual citations). Note that once a publication increases the CI of its author with an increment c, it contributes repeatedly to the utility of the author in each subsequent year, until the publication is superseded. If the paper is not superseded for k years, the total utility gained by its author is $k * c$.

The CBR and SelectivitY mechanism One explored scenario inspired from typical conferences is the Common Blind Reviewing (CBR) mechanism. CBR assumes double blind reviewing, where each paper is reviewed by M reviewers, and where each reviewer submits a comment and a score. Reviewers are allocated based on a bidding procedure that results in a high probability that papers are reviewed by researchers working on related topics. Any paper with a negative score is rejected. Papers with the average score above a given threshold are published as full papers and the remaining ones are published as posters.

We will also exemplify the use of the framework to analyze the mechanism for PR Games suggested in [3], that we call SelectivitY (SY). The decision policy for orally presented and poster articles is as for CBR, and it specifies that all other articles are registered as technical reports by the conference (except if authors decide to withdraw erroneous submissions). Comments and author answers are registered as technical reports by the conference[1] (implying that the peer-reviewing is not blind).

Experimentation with PR Games models can help evaluate the trade-offs between advantages and drawbacks of these mechanisms. For SY:

- The fact that all rejected submissions are registered as technical reports adds worth from the good papers that are rejected and reduces worth due to wrong papers that are registered.
- The worth lost due to wrong papers that are registered as technical reports is partly recovered by allowing authors to withdraw them, and also publishing the comments of the reviewers.
- The worth lost due to wrong comments is partly recovered by also publishing author answers.

4 Funding Based on Counting

Let us consider the case where the funding agency does not evaluate the citation influence according to the given formula, but simply rewards author i based on the number and venue of his publications:

$$R = n_o^i * w_o + n_p^i * w_p + n_t^i * w_t$$

where n_o^i is the number of his orally presented articles, n_p^i is the number of his posters, and n_t^i is the number of his technical reports. In this situation, the concept of paper superseding is no longer relevant, since papers are counted even if they are no

[1] or are digitally signed

longer cited. We will call these to be Trusted Peer-Reviewing Games (TPR Games).

For a two-players conference, with one submission each, and {accept,reject} decisions, the payoff matrix is:

	accept i's	reject i's
accept j's	1,1	0,1
reject j's	1,0	0,0

and the player's rational strategy for a one-shot version is to randomly make a decision.

For a repeated version of the game, a player can use forgiving Tit-for-Tat to force his colleague into not rejecting his paper. With forgiving Tit-for-Tat, rational players will converge into repeatedly play *accept*.

Paper Acceptance Thresholds This method for deciding rewards makes it rational for reviewers to accept all submissions. However, conferences that publish all submissions are given lower weights by funding agencies, which may compute weights $w_o, w_{,p}, w_t$ based on acceptance rates. Let us now assume that the conference (to remain relevant to the funding agency) puts a threshold on the ratio of accepted papers. The CBR reviewing mechanism modified to reject all the papers below the threshold (rather than papers with a negative review) is called CBRz.

If the conference decides to always accept for publication of a given percentage of the number of submissions, then the situation becomes a *zero-sum game*. Namely, the sum of utilities is given by the number of accepted papers, which is a constant.

In the new version we assume that the actions available to players are not {*accept,reject*} but the scores {*low, high*}. Let us assume that the conference accepts only 50% of the submissions. If a paper is scored higher than the other, only the higher scored paper is published. In the case of ties we will assume that the paper to be published is selected randomly. In this situation, the payoff matrix (with expected utilities) for 2 players is:

	high i's	low i's
high j's	1,1	0,2
low j's	2,0	1,1

Note that in one-shot versions of this game the rational action is to give a low score, since that has a higher payoff for each given action of the other player. For repeated versions of this two-payer game, the rational action remains the same as for one-shot games, due to the fact that the game is zero-sum.

Multiple players and Hits-for-Tat With n-players, even if the game is still a zero-sum game, it is not pair-wise a zero-sum game. For one-shot games and with blind review, the rational strategy remains to score low reviewed papers (which brings an expected utility of $\frac{1}{n-1} = \frac{n/2}{n-1} - \frac{n/2-1}{n-1}$ out of the value of a publication).

With non-blind review, forgiving Tit-for-Tat may seem an interesting strategy. The probability of having one's paper reviewed next round by the author of a paper scored low in the current round is asymptotically smaller than $\frac{1}{n-1}$ (due to the delay which allows the player to gain one reward in the intermediary year). Therefore with Tit-for-Tat the rational strategy remains to score *low*.

A promising strategy of player i for this scenario (that we call *Hits-for-Tat*) is a strategy where:

- k (two or more) submissions of j are scored low by i for each submission of i scored low by j.

Such a strategy brings a future expected penalty of $\frac{k}{n-1}$ from Tit-for-Tat for scoring low. With high values of k, the rational action when reviewing submissions of a Hits-for-Tat players is to score high. Therefore truthful reviewing is not in Nash equilibrium with a Hits-for-Tat player, either. Reported experiments prove this statement experimentally.

Discussion We note that with rational players, non-blind review (even with SY's strategy for publications) seems to not improve the social value of games where funding agency rewards are unrelated to the worth of the publications. Instead, the efficient strategies just shift from scoring low to new non-truthful reviews, such as scoring high with Hits-for-Tat. Other new non-truthful strategies can emerge, such as submitting many worthless articles to increase the probability for the publication of a submission.

5 Evaluating PR Games

We use simulations for experimentally evaluating the magnitude of the impact of different player strategies given the considered mechanisms (CBR, CBRz, and SY). We make this evaluation by computing the impact on one own's CI if the researcher writes wrong reviews for articles superseding his work. Some simulations assume that all researchers submit and review truthfully except for one reviewer. The reviewing strategies for the selected reviewer that we evaluate are:

(i) truthful reviewing,

(ii) truthful reviewing except for papers superseding one's work, which are rejected,

(iii) random reviewing except for papers superseding one's work, which are rejected.

(iv) giving the opposite possible score to all papers (reject good papers and accept poor papers).

(v) giving the lowest possible score to all papers.

For experimenting with funding based on CI, we generate 100 random research communities with 20 researchers over 20 conferences. While some researchers are more productive than others, for our experiments all participants are considered equally expert and inventive (since we only need to prove the **existance** of deviations from equilibria). In the generated problem instances, researchers get ideas for articles with a Uniform distribution at an average of 2 articles per year, and a worth that is uniformly random in [-10, 10]. Each paper is superseded each year with a probability of 1/5.

We report experiments for the weigths $w_0 = 0.5$, $w_p = 0.3$, and $w_t = 0.1$. Misclassifying comments have a negative worth (since the negative worth is canceled by answers, if published, it will be accounted only as bad reputation for the reviewer). The worth of a misclassifying comment is given by the displacement between the corresponding value V and the real worth of the paper (and it is always non-positive). An author's answer to a negatively misclassifying (i.e. assumed slanderous) comment has as worth the absolute value of the worth of the misclassifying comment (removing the negative effect of the reviewing comments on global worth and on the CI of the author). Otherwise, the answer has worth zero.

We analyze the following cases:

(a) all reviewers review truthfully

(b) all reviewers review truthfully, except for one reviewer who rejects articles superseding his work but reviews *truthfully* submissions not superseding his work

(c) all reviewers review truthfully, except for one reviewer who rejects articles superseding his work and reviews *randomly* submissions not superseding his work

(d) all reviewers review *truthfully* submissions not superseding their work and reject the other submissions

(e) all reviewers review *randomly* submissions not superseding their work and reject the other submissions

The average social value obtained for these cases is summarized in the following table:

	(a)	(b)	(c)	(d)	(e)
Worth(CBR)	967.8	931	878	219.77	21.38
Worth(SY)	967.8.8	939	896	408.68	224.07

The table shows us that the goal of the funding agency (social value) is maximized with truthful reviewing (a), and is reduced by cheating strategies. In SY with cases (b)-(e), even if all worthy papers are published, the total worth is reduced compared to the case (a) because many are published only as technical reports (being weighted with w_t to account for the low visibility).

Experiments with CBR confirm that the lying reviewer gains from under-evaluating papers superseding his work. To evaluate the equilibrium of truthful reviewing when the previously mentioned strategies are available, we select researcher 1 to perform non-truthful reviews. In average, researcher 1 earns 191.12 CI units by truthful reviewing when all other reviewers review truthfully (a). But, with CBR, he earns 493.33 by rejecting papers superseding his work, independently of whether he reviews the remaining papers truthfully (b) or randomly (c). However, with the SY mechanism, he will earn only 97,73 if he follows (b), and -131,1 if he follows (c).

	(a)	(b)	(c)	(d)	(e)
1's CI (CBR)	191.12	493.33	493.33	181.2	15.4
1's CI (SY)	191.12	97.73	-131.1	-0.81	-289.11

The experiments show the extent of the implications of the use of different strategies with CBR and SY. This confirms that truthful reviewing is not in Nash equilibrium when CBR is used, but (given our assumptions and given the possible strategies analyzed there) it is in Nash equilibrium when SY is used. Moreover, the social value is increased with SY versus CBR.

For the case of funding based on counting the number of publications, the experiments are based on CBRz. We simulate a community of 100 researchers. Each paper is reviewed by four researchers. We assume all reviewers review truthfully, except for one who reviews based on strategies (i), (iv), or (v). The papers are placed in descending order and the top (25%) are selected for the conference.

Using strategy (iv) a simulation was run 500,000 times and showed to have an advantage for the test subject. On average, the test subject improved their position by 9. The test subject was able to change his paper status from rejected to accepted 55,956 times (or 11.19%). For the cases when his cheating strategy made a change the social value of accepted papers changed from 11,050,688 to 10,858,837. This means that on average the test subject decreased the social value of this subset by a utility of 3.43 (0.38 overall).

Using strategy (v) the test subject was in average able to change his paper's position by 10 and its status from rejected to accepted 79,727 times (or 15.95%). For the cases when his cheating strategy made a change the social value of accepted papers changed from 15,627,432 to 15,119,179 (a change of 6.37 in utility, 1.01 overall). These experiments confirm that truthful reviewing is not in Nash equilibrium with CBRz.

6 Conclusions

We propose a new family of games, based on a simplified version of peer-reviewing. The new family of games can help to better study the mechanisms used for peer-reviewing and can help in the design of better mechanisms. We use simulations of the new game to prove that truthful reviewing is not in Nash equilibrium for a simplified version of the common blind reviewing mechanism (CBR), even without coalitions. We also study the mechanism suggested in [3], SelectivitY (SY), shown by simulations and by theory to provide a Nash equilibrium for truthful reviewing under the described assumptions and under the set of considered strategies. The main contribution of the section in this article that studies SY is that it exemplifies how game-theoretic models of conferences can help in designing better mechanisms, and it motivates the further study of such models.

We analyze four types of PR Games and find that, with the considered types of player strategies:

- for a common blind reviewing mechanism, CBR, a rational strategy consists of rejecting papers superseding reviewer's work and of producing random evaluations for the other papers
- for non-blind reviewing with paper rejections, an efficient strategy consists of an adapted forgiving Tit-for-Tat
- for non-blind reviewing that records all the submissions as either regular publications or technical reports (SY), with the studied set of available strategies, truthful reviewing was in Nash equilibrium and it maximizes the social value.
- for double blind reviewing and a fixed percentage of submission accepted (CBRz), truthful reviewing is also shown to not be in equilibria, and yields a lower social value.

References

[1] R. Axelrod. *The Evolution of Cooperation*. Perseus Books Group, 2006.

[2] V. Conitzer and T. Sandholm. Complexity of manipulating elections with few candidates. In *AAAI*, pages 314–319, 2002.

[3] J. Crowcroft, S. Keshav, and N. McKeown. Scaling the academic publication process to internet scale. *Communications of the ACM*, 52(1):27–30, Jan 2009.

[4] M. Dresher. *The Mathematics of Games of Strategy: Theory and Applications*. Prentice-Hall, 1961.

[5] M. Flood. Some experimental games. Technical Report Research memorandum RM-789, RAND Corporation, Santa Monica, CA, 1952.

[6] A. Gibbard. Manipulation of voting schemes. *Econometrica*, 41:587–602, 1973.

[7] J. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.

[8] A. Rapoport. *Decision Theory and Decision Behavior*. Kluwer, 1989.

[9] M. Silaghi and M. Yokoo. The peer-reviewing game. Technical Report CS-2008-03, Florida Institute of Technology, November 2008.

[10] W. Vickrey. Counterspeculation, auctions and competitive sealed tenders. *Journal of Finance*, 16:8–37, 1961.