11-8-2008

The Peer-Reviewing Game

Marius C. Silaghi

Makoto Yokoo

The Peer-Reviewing Game

M.Silaghi and M.Yokoo
Florida Tech and Kyushu University
FIT Technical Report
CS-2008-03

ABSTRACT

We model a family of peer-reviewing processes as game-theoretic problems. The model helps to understand elements of existing peer-reviewing procedures, and to predict the impact of new mechanisms. The peer-reviewing for evaluation of scientific results submitted to conferences involves complex decision processes of independent participants. Significant tax-payer money is spent by governments for the advancement of science, and such governmental decisions are often based on outcomes of peer-reviewing in scientific conferences. Here we propose to analyze and design improved mechanisms for conference peer-reviewing, based on game-theoretic approaches. Real world conference peer-reviewing processes are overly complex and here we define and analyze a simplified (toy) version, called *the Peer-Reviewing Game*. While our toy version may have significant assumptions, it provides an interesting game and a first step towards formalizing and understanding the real world problem.

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. In this work the conference chair is assumed to be a trusted party, enforcing policies agreed by the funding agency and making publication decisions based on the recommendation of the reviewers. We uncover relations between Peer-Reviewing games and Prisoner's Dilemma games. Examples of mechanisms are described and analyzed both theoretically and 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 [15], voting processes [4, 7], and in the Prisoner's Dilemma [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. The approach here uses simplifications to model some elements (such as citations, reviewer assignment, program chair selection, reviewer expertize, and funding agency's decisions), in order to avoid getting entangled in excessively many details.

Cite as: The Conference Peer-Reviewing Game and Prisoner's Dilemma, Author(s), *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX-XXX.

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. Features at fashion include:

- blind reviewing: namely assuring the secrecy of the reviewers' names, and sometimes (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.

While we here address some of these features, we leave the modeling of other features for future research.

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. 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. For example, the community *gains utility* from:

- an article with an original scientific contribution.
- a well written tutorial, even if it disseminates only existing knowledge,

Some cases when the society *loses utility* are:

- when an erroneous article is published,
- when an irrelevant article is orally presented and takes valuable time from the audience,
- when lack of appropriate citations in published papers frustrates other researchers and makes them abandon research,

- due to not publishing an article which by the previous definition would have increased utility.

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. Public confirmation that a review was incorrect happens in the two-players version of a peer-reviewing game (see more in Section 4). Public confirmation of the incorrectness of a review can also happen in conferences without blind review.

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

Since we assume that the funding agency cannot directly measure the utility (worth) of the work, it distributes funding based on a *citation influence* computed based on the number of citations to papers presented in conferences. Therefore, cheating strategies by researchers can be motivated by a desire to increase their citation influence (funding):

- author claiming original contributions by not providing proper citation (plagiarism),
- reviewer 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 the next section we introduce background about game theory and Prisoner's Dilemma, found to be related to our problem. After introducing in Section 3 the formal definitions of the main involved concepts, Section 4 analyzes simple Peer-Reviewing Game problems with non-blind review. The next section discusses some of the implications of blind reviewing. In Section 6 we describe the features taken into account in a slightly more complex version of the peer-reviewing problem. Subsequently we introduce the formal framework, the Peer-Reviewing Game (PR Games), used here to model such problems. Further, a commonly used mechanism for playing PR Games is formalized, as well as a newly proposed mechanism. In Section 12 we evaluate 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

We find that our problem has elements that are related to the classic Prisoner's Dilemma.

Prisoner's Dilemma.

The merit of stressing that decision making can involve contradictory drives to conflict and cooperation is attributed to [1] who analyzes the repeated Prisoner's Dilemma (PD) problem. According to [9], more than a thousands articles were written on this topic. In one version of the Prisoner's Dilemma [6, 5], 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) [2]. 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* [13]. 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 player tries to avoid on 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 a famous strategy that performs extremely well (for 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. An example of irrational player is the superrational player. According to [8], a *superrational player* is one that plays according to a strategy obtained by maximizing utility under the assumption that all players reason in the same way. The superrational player would try to play against a player of an unknown type using a probabilistic approach based on what the superrational player expects from the second player, trying to use the same strategy.

Prisoner's Dilemma has been applied to many other problems (with slightly different values in the payoff matrix, but with similar issues), such as *Friend or Foe* [11] or *Hawk and Dove* [12]. Game theory has been applied to many other problems, such as bidding in auctions and voting [15, 4].

A set of strategies in a game are in Nash equilibrium [10] if no player can do better by changing his strategy while the strategies of the other players remain fixed.

3. DEFINITIONS OF MAIN CONCEPTS

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

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. For simplicity we assume that all reviewers in the community are experts.

DEFINITION 2 (SOCIAL VALUE). *The social value consists of the sum of the worth of the papers published in the community.*

DEFINITION 3 (SUPERSEDED PAPER). *A paper P_1 is superseded by a paper P_2 if P_1 no longer contributes to the influence of its author after the publication of P_2.*

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 types: regular, poster, technical report. The corresponding weights are w_o, w_p and w_t.*

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' CI.

DEFINITION 5 (FUNDING AGENCY). *The* funding agency *is the entity that yearly pays each researcher an amount equal with their total CI at that moment.*

While the funding agency cannot access the worth of a paper directly, 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 concept of blind review is different in different communities. We use the following flexible definition.

DEFINITION 6 (BLIND REVIEW). *A peer-reviewing mechanism is said to be based on* blind review *if reviewers learn the name of the authors of the submitted papers, but authors do not learn the name of the reviewers of their papers.*

To avoid confusions, we also define double-blind reviewing.

DEFINITION 7 (DOUBLE-BLIND REVIEW). *A peer-reviewing mechanism is said to be based on* double-blind review *if neither do reviewers learn the name of the authors of the submitted papers, nor do authors learn the name of the reviewers of their papers.*

DEFINITION 8 (NON-BLIND REVIEW). *A peer-reviewing mechanism is said to be based on* non-blind review *if reviewers learn the name of the authors of the submitted papers and authors learn the name of the reviewers of their papers.*

In this article we have to stick to the above definition for blind review, even if we are aware that many readers typically use the name *blind review* for what we define as *double-blind review*.

4. TWO-PLAYER PR-GAME

Before considering peer-reviewing games involving many researchers, we will illustrate the basic problem on a case with two researchers, named researcher 1 and researcher 2. The researchers participate as authors and reviewers in a suite of K conferences (defining an iterative game). K may be finite or infinite. Since we assume to have only 2 researchers, each submitted article can receive a single peer review. Namely, in each conference, each researcher submits an article that will be reviewed by the other researcher. This means that the reviewing cannot be blind. Each article has a worth in the range $[-MAX, MAX]$. First let us assume that an article can be either rejected, or accepted as a regular orally presented paper.[1] Each published article increases the *citation influence* (CI) of its author. For simplicity we assume that this increment is given by the worth of the paper, weighted with a constant, w_o. If $CI_{i,k}$ is the citation influence of researcher i immediately after the conference k, where he publishes a regular paper with worth ${W^{pub}}_{i,k}$, then:

$$CI_{i,k} = CI_{i,k-1} + {W^{pub}}_{i,k} * w_o$$

The utility $u_{i,k}$ of each researcher i after conference k increases with an increment given by its current CI.

$$u_{i,k} = u_{i,k-1} + CI_{i,k}$$

We also assume that an article may supersede older articles, removing their citation influence. We therefore obtain:

$$CI_{i,k} = CI_{i,k-1} + ({W^{pub}}_{i,k} - {W^{sup}}_{i,k}) * w_o$$

where ${W^{sup}}_{i,k}$ is the worth of older publications of i superseded at conference k.

If both researchers' papers are rejected, each gets a utility based on his old CI which is unchanged. If researcher j's paper is rejected while researcher i's paper is accepted, i wins the contribution to his CI of his paper's worth W_i projected over the remaining $K-k$ years, namely $Wi = W_i * w_o * (K-k)$. Meanwhile, j loses the contribution to his CI of his papers with worth W'_j superseded by i's new publication, projected on the remaining k years, $Wj' = W'_j * w_o * (K-k)$. Wj and Wi' are defined symmetrically, specifying the payoffs when j's paper is accepted and i's paper is rejected. If both papers are accepted, then player i wins $Wi - Wi'$ (the combination of the additional CI due to the new paper and of the CI lost due to superseding by the other's paper).

Assuming the articles have a positive worth (not creating bad reputation if published), then the corresponding payoff matrix in one iteration is:

	accept i's	reject i's
accept j's	$Wi - Wi', Wj - Wj'$	$-Wi', Wj$
reject j's	$Wi, -Wj'$	$0, 0$

The first item in each cell specifies the payoff of i and the second item specifies the payoff of j. Due to symmetry assumptions, in some of the following payoff matrices we only specify the item for the payoff of i.

The rational strategy with positive Wi' and Wj' is reject. One interesting result is that in this simple model of the peer-reviewing game, there is a potential inverse relation between the quality of the paper and the incentive of the

[1] Not yet considering posters, technical reports, etc.

reviewer to accept the paper. Namely, in real world a better paper can be expected with higher probability to supersede a paper of the reviewer (leading to a higher Wi' or Wj'). In PR Games this encourages the reviewer to reject the paper.

If researcher j's paper does not supersede researcher i's work then $Wi' = 0$ and i's payoff matrix is:

	accept i's	reject i's
accept j's	Wi	0
reject j's	Wi	0

showing that the player i can be indifferent about the publication of player j's article (e.g., can reject it if it is so required by a Tit-for-Tat strategy).

Accepting an erroneous paper in the 2-player case implicitly reveals the error of the review. The negative worth of the erroneous review is considered to be equal with the negative worth of the erroneous paper. If player j's paper has a negative worth, then we have to take into account the *bad reputation* it can produce for its reviewer, and i's payoff matrix is:

	accept i's	reject i's
accept j's	$Wi - Wi' - \|Wj\|$	$-Wi' - \|Wj\|$
reject j's	Wi	0

where Wj is the contribution to CI of player j's paper, with absolute value $|Wj|$. In a one-shot game, a rational player i will reject player j's article.

In the case where Wi, Wi', and Wj are all positive, and Wi' is lower than Wi, then i's payoff matrix resembles the one in Prisoner's Dilemma. For example, if $Wi' = Wi/2$ then:

	accept i's	reject i's
accept j's	$Wi/2$	$-Wi/2$
reject j's	Wi	0

This shows the corresponding scenario of the PR Game to be similar to Prisoner's Dilemma [6]. We infer that an adapted version of forgiving Tit-for-Tat strategy may perform well in an infinite iteration of such scenarios. Namely, in this version a researcher retaliates only if a paper with positive worth is rejected (not if a paper with negative worth is rejected). The adapted forgiving Tit-for-Tat strategy implements a basic *reputation system* [3].

We can assume that in the case with only two researchers there is no plagiarizing. This is because any plagiarized paper (paper originally authored by the other researcher), would have to be reviewed by its original author (who can reject it).

A Generalization.

We can generalize this result to a version of the n-players PR Games, namely the version with non-blind reviewing and with accepted/rejected decisions.

For repeated n-player PR Games with an infinite number of iterations, non-blind review, and where the publication decision can consist of either *regular papers* or *rejection*, we can anticipate that the adapted forgiving Tit-for-Tat strategy will be a reasonable robust policy. Namely, each PR Game of this type is equivalent to a set of parallel two-player PR Games, one between each pair of researchers. Therefore a good strategy of the n-player version is a composition of the strategies of the two-player versions, a set of adapted forgiving Tit-for-Tat strategies towards each peer.

5. BLIND REVIEW

The two-player version of the peer-reviewing game fails to capture important aspects of real world peer-reviewing processes. In particular, blind review makes Tit-for-Tat strategies impossible in real mechanisms, and also reduces opportunities for implementing reputation systems. Moreover, it cannot model strategies based on plagiarism. To enable the modeling of these phenomena, we need to consider versions with at least three researchers. For example, with three researchers and a single review per submission, there is an uncertainty as to who was the reviewer (the reason why Tit-for-Tat strategies do not apply).

In some conferences with several reviewers per paper, the reviewers see each other's names and reviews, and some limited *bad reputation* can be developed (limited because reviewers are typically expected to not make public another reviewer's bad performance). However, in most conferences reviewers do not see each other's name and therefore, even if they can see each other's review, *bad reputation* cannot be developed. Program chairs do see the names of reviewers but they are not expected to read the papers and to know their real worth (and may not be experts on their topic), so they may not be expected to participate in a *bad reputation* mechanism.

We therefore assume in this simplified model that there is no *bad reputation* mechanism with blind reviews. For each pair of two researchers i and j, the blind review causes i's payoff matrix to not have a *bad reputation* component for accepting j's poor papers, obtaining always:

	accept i's	reject i's
accept j's	$Wi - Wi'$	$-Wi'$
reject j's	Wi	0

If we assume that reading an article may not directly benefit a reviewer, then we can state:

LEMMA 1. *With blind review and no bad reputation mechanism, a rational strategy for i is to reject j's paper whenever $Wi'>0$. If $Wi'=0$ then i's payoff is independent of the decision of i, and writing a* random review *is a rational strategy for i.*

PROOF. The case with $Wi' > 0$ follows directly from the payoff matrix. If one would assume that reviewing takes time (and utility) from the reviewer, than the reviewer has an incentive to not even read articles suspected to not supersede their own work, but to just propose randomly a decision. □

Plagiarizing.

We first formally define the plagiarizing strategies for Peer-Reviewing Games.

DEFINITION 9 (PLAGIARIZATION). *A plagiarizing player is a player that submits articles that he had reviewed and rejected in an earlier conference.*

With random reviewer assignment, a plagiarizer has a 50% probability to succeed getting another reviewer than the original author of the article. We will assume that, if the reviewer is the original author of the article, then the plagiarizing submission will be rejected. With typical conferencing mechanisms, plagiarizers do not get a bad reputation. This is because peers have no way of verifying whether the older paper was used as a base of the plagiarized one (assuming the plagiarizer rephrases it, as recommended by lawyers for avoiding copyright infringements [14]).[2] Even the original author cannot know whether the new paper is a plagiarizing paper or a rare coincidence, since blind review preclude him from learning the identity of his reviewers. The original authors may not be legally allowed to contribute to a bad reputation of the new authors, because of their lack of proof. Moreover, a Tit-for-Tat approach to plagiarizing is not possible because blind review precludes the original author from knowing when they are reviewing articles of a suspected plagiarizer.

We further introduce the common blind review mechanism (CBR). We experimentally verify that truthful reviewing with CBR is not in Nash equilibrium (assuming all the other reviewers are truthful), e.g., since rational reviewers have incentives to both plagiarize, and to reject worthy articles superseding their work.

In the next section we define more formally the general PR Game problem, used for experimental results verifying the extent of the impact of blind review on the worth produced by the PR Game and for verifying the lack of equilibrium for truthful reviewing.

6. PEER-REVIEWING GAMES

Here we consider that a Peer-Reviewing Game (PR Game) consists of a community of N researchers, $\{1, ..., N\}$, participating in a series of conferences $\{1, ..., K\}$, e.g., where conference k is scheduled in the k^{th} year. We assume that each researcher i produces between the deadlines of the conferences $k-1$ and k a number of $O_{k,i}$ papers, namely the entries (k, i, t) in a matrix $P[K, N, T]$, where $t \in [1..O_{k,i}]$, and T is the maximum value for $O_{k,i}$. An author can submit to conference k these $O_{k,i}$ new papers , older papers rejected at previous conferences, and/or other unoriginal papers. In PR Games, each submitted paper will be peer-reviewed by M researchers selected using a *reviewer policy* P_R (e.g., random assignment, paper bidding). For simplicity we assume that each paper has a single author.

The utility of publishing each paper is given by the *worth* of the paper. Each paper p has a worth $W_p \in [-MAX, MAX]$. The problem description will store the worth of each paper as an entry in a matrix $W[K, N, T]$ whose dimensions are as for the matrix P.

Each reviewer m has to assign to each paper p an estimated value $V_{m,p}$ and a comment $C_{m,p}$, to which the author gives the answer $A_{m,p}$. Comments and answers contain information and therefore they have their own worth. A *decision policy* P_D uses the comments and the set of values assigned by reviewers to decide for each paper and reviewer comment whether it is published as regular paper, as poster, as technical report, or simply denied.

[2]Also original authors may not be even able to prove their priority since records of old submissions are not normally made available, and blind review precludes verifiers from learning that the plagiarizer has reviewed the original submission.

A paper is *superseded* by another paper when the second paper presents a better solution. This happens for a given set S of pairs of papers.

With the intent of maximizing the total worth of the articles in the community, a funding agency funds each researcher with an amount proportional with his *citations influence* (CI). Namely, we assume that a researcher i gets a citation influence CI given by $CI = v_o^i * w_o + v_p^i * w_p + v_t^i * w_t$ where v_o^i is the worth of his unsuperseded orally presented articles, v_p^i is the worth of his unsuperseded posters, and v_t^i is the worth of his unsuperseded technical reports. By w_o one denotes a weight representing the impact of oral presentations, by w_p the impact of posters and by w_t the impact of technical reports for the given conference.

A mediator, called program chair, enforces the policies for assigning reviewers and the policies for deciding the publication of individual papers based on reviewer recommendation. For simplicity the mediator is assumed to be honest.

DEFINITION 10 (RESEARCH COMMUNITY). *A research community is defined by a tuple* $\langle N, K, O, P, W, S, w_o, w_p, w_t \rangle$ *where*

- N *is the number of researchers (they act as both authors and reviewers),*
- K *is the number of conferences (we assume that there is one conference each year),*
- $O[K, N]$ *is a matrix where each element* $O_{k,i}$ *specifies the number of original papers prepared by author* i *prior to the deadline for conference* k*, and after the deadline for conference* $k-1$ *(if* $k>1$*),*
- $P[K, N, T]$ *is the matrix containing at most* $x = K*N*T$ *new papers, where* T *is the largest number of papers per author per conference time-frame, and where some entries are empty if the corresponding author produces less than* T *papers in that conference time-frame,*
- $W[K, N, T]$ *is the matrix of worths for these papers, the worth of each paper being known only to peer-reviewers, and to their author after reading the reviews,*
- S *is a set of pairs of articles,* (p_s, p)*, where* p *supersedes* p_s*, each paper being specified by a triplet* (k, i, t) *showing its position in the matrix* P*,*
- w_o*,* w_p*,and* w_t *are the impact weights of unsuperseded regular articles, poster articles, and technical reports, respectively, and are used in the computation of citation influences* CI *for researchers.*

Researchers are rewarded yearly with funds computed using their CI. No new researcher joins the community, but a researcher will leave the community with a probability Pr if reviewers misclassify one of his articles.

The mechanism design problem consists of defining rules for assigning reviewers and for deciding paper publications in order to maximize the social value (the worth of published papers). The PR Game family of games is defined as:

Definition 11 (PR Game). *A Peer-Reviewing Game (PR Game) is specified by:*

- *a research community* $\langle N, K, O, P, W, S, w_o, w_p, w_t \rangle$,
- *the number of reviewers assigned per paper,* M,
- *the policy for assigning reviewers,* $P_R : [1..x] \rightarrow [1..N]^M$, *for* x *submitted papers,*
- *the policy for deciding what to do with the paper,* $P_D : (\mathcal{V}^M, \mathcal{C}^M, \mathcal{A}^M) \rightarrow \{oral, poster, report, rejection\}$, *with* $\mathcal{V}, \mathcal{C}, \mathcal{A}$ *specifying the reviews, as defined below,*
- *the policy* $P_H : \emptyset \rightarrow \{report, discard\}$ *of how to handle reviewer comments and author answers, discarding them for blind review, or recording them as technical reports (e.g., with non-blind review).*

where a reviewer can assign to a paper a numerical evaluation from a set $\mathcal{V}$ *and a comment from a set* $\mathcal{C}$. *An author can answer each review with an answer from a set* $\mathcal{A}$. *Elements of* $\mathcal{C}$ *and* $\mathcal{A}$ *have worth given by functions*

$$P_C : ([-MAX, MAX] \times \mathcal{V} \times \mathcal{C} \times \mathcal{A}) \rightarrow [-2MAX, 2MAX]$$

and

$$P_A : ([-MAX, MAX] \times \mathcal{V} \times \mathcal{C} \times \mathcal{A}) \rightarrow [-2MAX, 2MAX].$$

For each review, these functions map the worth of the paper and the review into a worth of the comment, and of the answer, respectively. The worth of a comment has impact both on the CI of the reviewer and on the CI of the author of the submission. The worth of an answer only impacts the CI of it author (possibly correcting the influence of wrong reviewing comments).

Together, the submission, the reviewing comments, and the answers, have a total impact on the social value given by the actual worth of the submission.

Data used by considered mechanisms.

During the solution process, the reviewers are required to produce value assessments and reviewing comments to be used in the decision policy, and authors may submit answers. If T is the maximum number of papers submitted by an author to a conference, then:

- $V[K, N, T, M]$ is the matrix of assessed values attached to each paper by its reviewers, with values from a set $\mathcal{V}$,
- $C[K, N, T, M]$ is the set of comments submitted by reviewers, with values from $\mathcal{C}$,
- $A[K, N, T, M]$ is the set of answers to reviews by the authors, with values from $\mathcal{A}$,

Simplifications.

The main simplifications of the PR Game over typical conference peer-reviewing is that:

- we assume in experiments that the horizon (total number of conferences) is finite. However, the studied strategies were not adapted to let players exploit the known finite horizon [13].
- we assume that each paper has a single author.
- we assume that new researchers do not appear.
- we assume that authors accept to review any assignment.
- we assume that reviewers are rational (self-interested).
- we assume that all the reviewers are experts in the same field.
- we assume authors are rewarded according to the citation influence that we define.
- we assume authors submit only their new papers, only once, and only at one conference.
- we assume in experiments that reviewing comments do not propose improvements to the reviewed article, but in the real world peer-reviewing comments can be constructive and, if published, can add good reputation (and increase the CI) for the reviewer.

Mechanism design for PR Games.

The PR Game mechanism design problem consists of tuning the different parameters of the game in order to guarantee:

- a maximization of the funding-agency's targeted criteria of maximizing the total social value of the papers.
- a strategy-proof robustness from the point of view of the reviewers commenting and evaluating.

Some possible non-truthful strategies are:

- researchers can make coalitions.
- reviewers can under-evaluate an article superseding their work (or work of researchers in their coalitions).
- reviewers can wrongly evaluate articles as a result of a Tit-for-Tat strategy, or of a random review strategy.
- researchers could negotiate offline and exchange money to accept sets of papers maximizing the sum of CIs (mainly with non-blind review).
- reviewers can under-evaluate articles to be able to plagiarize them.
- researchers can submit plagiarized articles (i.e., without proper citations).

The total worth produced by the conference is computed by counting a single time duplicates due to plagiarism.

7. SAMPLE MECHANISMS

In this section we present two mechanisms for solving PR Games. One of them is a simplification of commonly used reviewing procedures for AAMAS. The second is a new reviewing procedure that we propose here. We are going to later present preliminary comparison of the two mechanisms based on randomly generated PR Games.

The simplified CBR mechanism for PR games.

The Classic Blind Review (CBR) reviewing is a version of a blind review mechanism (used in conferences such as AAMAS08). We will assume for simplicity that there are $M = 3$ reviewers per article. The P_R policy to select reviewers is based on a bidding phase. We will assume for simplicity that the authors of papers superseding or superseded by an article are assigned to it as reviewers with 50% probability.

In our experimental evaluation we make the approximating assumption that the policy P_D to decide the fate of submitted articles in CBR is based on rejecting any article p where a reviewer declares that $V_p \leq 0$, and selecting for oral presentation the rest of the articles p where no reviewer declares that $V_p \leq MAX/2$. The P_H specifies that no reviewer's comment or author's answer is published.

The SelectivitY mechanism for PR games.

We will also exemplify the use of the framework to analyze a new mechanism for PR Games, that we call SelectivitY (SY). In the SY mechanism, the parameters M, and P_R are as for CBR. The P_D decision policy for orally presented and poster articles is as for CBR, except that it specifies that all other articles are registered as technical reports by the conference (except if authors decide to withdraw erroneous submissions). P_H specifies that comments and author answers are registered as technical reports by the conference (implying that the peer-reviewing is not blind).

The SY mechanism has advantages and drawbacks, and experimentation with PR Games models can help evaluate the trade-offs between advantages and drawbacks.

- 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.

We assume for simplicity that author answers do not induce negative worth. PR Games may not yet model all influences on reviewing that come from the fact that the review is no longer blind (psychological, social, etc.), which are subject of further research.

8. TWO-PLAYER SY CASE

If the PR game is based on using the SY mechanism, the payoff matrix changes due to the penalties induced by the new type of bad reputation. Namely, while the non-blind reviewing mechanism studied in Section 4 allows for bad reputation only when erroneous papers are published, the SY mechanism introduces a bad reputation when worthy papers are rejected. So, if player j's submission is publication worthy, i's playoff matrix is:

	accept i's	reject i's
accept j's	$Wi - Wi'$	$-Wi'$
reject j's	$Wi - Wi' - \|Wj\|$	$-Wi' - \|Wj\|$

Here the papers are always published (as either regular paper or technical report), and the impacts Wi, Wj, Wi', have to be computed with the corresponding weights. Therefore the superseded work of i loses its contribution to i's CI for all alternatives (always subtracting Wi'). The penalty for a poor review rejecting j's submission is Wj and is subtracted only in that event. We see that with SY, independently of j's action, the rational behavior of the player i is to correctly review j's paper. The same holds when j's paper is not publication worthy, as seen in Section 4.

9. MORE ON SELECTIVITY

In order to increase the social value (total utility of published articles), new peer-reviewing mechanisms may have to be designed. We evaluate the impact from the introduction of a reputation system based on the strategy called SelectivitY (SY).

SY builds a reputation system based on non-blind review. In the experimented reputation system, each external expert can independently verify correctness of reviews and plagiarizing by having access to old submissions and to the reviews of each researcher. For this purpose, all submissions, reviews, and author answers to reviews are registered as technical reports. Now, the CI of a researcher misclassifying a submission in a review gets a penalty, proportional to the deviation of his review, weighted with a constant specifying the impact of technical reports, w_t. Similarly, authors of worthy articles that are rejected (i.e., registered as reports) get an increase in their CI , and plagiarizers get a penalty.

Researchers submitting erroneous papers are also penalized since their reputation suffers, given that the rejected submissions are now accessible as technical reports. We remind that authors learn the worth of their contribution only after seeing the reviews. To ensure that participation in the game is rational (players do not lose by participating), SY allows researchers to withdraw papers with negative worth.

Note that throughout this research we assume that there exists a mechanism which ensures that the worth of published articles, as well as bad reputation, will be mirrored in the CI computation of the funding agencies (e.g., through some kind of citation counting).

THEOREM 1. *SY is truth incentive for Peer-Reviewing Games, with respect to misclassification of articles and plagiarism.*

PROOF. Both deviations from truthful behavior only bring reduction to the CI (utility) of the researcher. □

The effect of SY on the social value as compared to the case of truthful reviewers and CBR is described in experiments.

10. EVALUATING PR GAMES

We use simulations for experimentally evaluating the magnitude of the impact of different player strategies given the considered mechanisms (CBR 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 assumes that all researchers submit and review truthfully except for one reviewer that under-evaluates the papers that supersede his work. The reviewing strategies 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.

Another simulation assumes that all researchers submit and review truthfully except for one reviewer that plagiarizes. Plagiarized authors leave the research community with a probability of 50% after each successful plagiarization.

We generate 100 random research communities with 20 researchers over 20 conferences and generated assuming that researchers get ideas for articles with a Uniform distribution at an average of 2 articles per year, and a worth that is uniformly random between -MAX and MAX, where MAX=10. Each paper is superseded each year with a probability of 1/5. We also use $Pr = 0$.

Numeric assumptions in experiments.

We assume that $w_0 = 4$, $w_p = 2$, and $w_t = 1$. In our experiments we assume P_C and P_A to be as follows. 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.

All reviewers are assumed to be experts, and researchers are assumed to leave the community with a probability $Pr = 1/4$ if their work is negatively misclassified (or plagiarized).

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 total worth obtained for these cases is summarized in the following table:

	(a)	(b)	(c)	(d)	(e)
Worth(CBR)	834.8	803.1	755.83	219.77	17.05
Worth(SY)	834.8	804.64	759.13	247.35	53.28

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, even if all worthy papers are published, the total worth is reduced 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.

11. FUTURE WORK FOR PR GAMES

We made simplifications, such as assuming that Program Chairs (PCs) are honest mediators. In future work we plan to evaluate the case where PCs can be part of coalitions with other researchers. He can also act untruthfully, e.g., for articles superseding his own work.

In some conferences, articles are restricted to contain a limited amount of citations or of space. Rational reviewers and the PC may prefer articles that cite them. Authors have to decide which citations to select in order to maximize the satisfaction of probable reviewers. Truthful schemes need further research in future work.

Expertize of researchers may not be equal (but quantifiable with a value), and worth of articles and reviews may depend on this expertize. Funding agencies may use different mechanisms for funding decisions.

12. SIMPLER FUNDING AGENCY

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

If the conference decides to always accept for publication 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*}. We also 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) 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).

However, 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. 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.

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.

In future research we hope to identify reputation mechanisms whereby reviewers make a coalition to score low submissions of non-truthful reviewers, without getting bad reputation (which SY does not offer). The availability of such a mechanism may bring equilibrium for a strategy, which we call *truthful reviewing with k punishing exceptions*:

- the coalition agreement is that each player i has to score truthfully each submission of any player j, if player j is not under collective punishment.
- if a player is under collective punishment, then the coalition agreement is that his current submission has to be scored low.
- a player j is under collective punishment if and only if player j scored in disagreement with the coalition agreement a submission in a previous conference, and has not been yet punished for it with low scores for k good papers.

THEOREM 2. *The* truthful reviewing with k punishing exceptions *strategy is in Nash equilibrium if* $k \geq 1$.

13. 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 show that truthful reviewing is not in Nash equilibrium for a simplified version of the common blind reviewing mechanism (CBR), even without coalitions. We also describe a mechanism, 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. Given the simplifications and assumptions that we use to get a simple model, we must foresee that the proposed SY mechanism may not perform similarly well when used in real world peer-reviewing. 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 three 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 and 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 is in Nash equilibrium and it maximizes the worth of the conference.

We also investigate the case where funding agencies reward researchers based on the number of publications rather than based on the number of citations (TPR Games), and find that the equilibria change dramatically. We propose a different strategy that is in Nash equilibrium in such games and which provides incentives for truthful reviewing.

Given the disastrous efficiency predicted by our model for the common blind peer-reviewing mechanism, the reader may wonder why, in the real world, related peer-reviewing mechanisms allowed the technology to develop so well. The perceived difference can be explained by some of our simplifications, and mainly the fact that real researchers may often be altruistic (superrational) rather than rational. However, the results signal that changing reviewing mechanism may bring improvements. We plan to perform additional research (with less simplifications and more strategies) to predict with better accuracy how new mechanisms would perform with real world conferences.

14. REFERENCES

[1] R. Aumann. Acceptable points in general cooperative n-person games. *Contributions to the Theory of Games*, IV (Annals of Mathematics Study)(40):287–324, 1959.

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

[3] R. Boyd. Mistakes allow evolutionary stability in the repeated prisoner's dilemma game. *Journal of Theoretical Biology*, (136):47–56, 1989.

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

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

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

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

[8] D. R. Hofstadter. *Metamagical Themas: Questing For The Essence Of Mind And Pattern*. Basic Books, 1996.

[9] S. Kuhn. Prisoner's dilemma. `plato.stanford.edu/entries/prisoner-dilemma`, 2007.

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

[11] F. Oberholzer-Gee, J. Waldfogel, and M. W. White. Social learning and coordination in high-stakes games: Evidence from 'friend or foe'. Social Science Research Network, 2003.

[12] Osborne and Rubenstein. *A Course in Game Theory*. MIT Press, 1994.

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

[14] D. Stewart. Patent infringement–making, not losing money. FIT Seminars, 2008.

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