Incentivizing anonymous “peer-to-peer” reviews

Parv Venkitasubramaniam
Electrical and Computer Engineering
Cornell University
Ithaca, NY 14850
Email: pv45@cornell.edu

Anant Sahai
Electrical Engineering and Computer Science
University of California, Berkeley
Berkeley, CA 94720
Email: sahai@eecs.berkeley.edu

***Abstract*—The review cycle for papers takes way too long in many disciplines. The problem is that while authors want to have their own papers reviewed fast, that are often unwilling to review the papers of others in a timely manner. This paper explores what would be required to incentivize fast reviews using a public reputation/scoring system that exploits the fact that the referees are drawn from the same pool as paper authors. The challenge in maintaining a public reputation system is to ensure that the identity of referees remain as anonymous as possible. A model is proposed in this work, wherein authors have an incentive to commit to reviewing papers and are rewarded for meeting this commitment in a manner that prioritizes their own papers for reviews. This ensures stability (bounded reviewing delays) for all fair contributors while freeloaders face a potentially unstable system. A naive implementation of the scoring system, however, leaks information that would allow authors to infer the likely identities of their referees. A distortion to the observed public score process is then studied, which is shown to enhance anonymity while preserving the incentives for timely refereeing.**

I. INTRODUCTION

With possible competition from the lack of good parking spaces[1], the number one complaint of many researchers is that papers take an unreasonably long time to get fairly reviewed. Arguably, the only reason that researchers do not complain more vocally about this is that each of us has the secret shame of a few unreviewed papers sitting in our offices that we just have not gotten around to yet. Herein lies the seeming paradox: while we cannot build parking lots for ourselves, the community of researchers is itself responsible for the slow peer review of its own papers. This article proposes and analyzes a reputation based system that could expedite this process by aligning the incentives of reviewers and the community.

In the next subsection, we touch briefly on related work studying peer review and general peer-to-peer systems. Section I-B sketches the features of our proposed system to incentivize timely peer-review. With the proposal in place, a simplified mathematical model is described in Section II. The analysis based on the model then proceeds in two stages. First, Section III shows how the proposed system can meet the desired objectives of fairness to good scholarly citizens by assuring them of timely reviews. Second, Section IV uses a closely related model to argue how referee anonymity can be preserved despite having public scores. Finally, Section V concludes this article with some comments on the tension between the two objectives.

A. Related Work

At first glance, the problem of peer review seems to be a problem of incentive misalignment in the classic “tragedy-of-the-commons” mold [2] — the pool of referee time is a limited public resource (like a common grazing area) and thus people inject more papers (cattle) than the system can stably serve leading to delays. Peer review has been the subject of numerous studies, but the space limitations restrict us from doing complete justice to the literature here[2]. (See [3]–[5] for a survey of peer review in general.) The slowness of peer-review is explicitly considered in the literature[3] with some even arguing that such delays serve as deterrents to oversubmission (performing the role of a flow control signal as in TCP) in the absence of any other credible deterrent for such submissions [10]. However, the community has also understood there is an undersupply of reviewing time and explicit pricing mechanisms have been proposed using real money[4] [11], [12] as well as using non-cash tokens that are earned by reviewing papers and spent by submitting papers [10], [13]–[15]. Peer-pressure and status-consciousness resulting from public reputations have also been discussed [10].

---

[1] Clark Kerr’s iconic 1957 remark [1] was that the alumni seem to care mainly about athletics, the students mainly about sex, and the faculty mainly about parking. To be realistic, the list of commonly aired complaints also includes (in no particular order) how hard it is to get funding nowadays, how we are all paid far too little in relation to our true worth, how students these days are just not as strong as they were in the good old days, and how hard it is to get a good job for our students/postdocs. This paper will have nothing new to say about any of these other topics and the interested reader is referred to any casual gathering of more than three faculty members.

[2] As many point out, peer-review is surprisingly unstudied given how central it is to our science-driven society. We suspect that a case can be made that peer-review should join “law making” and “sausage making” on the list of things that are best appreciated at a distance and should not be studied too closely lest we and the public loose all faith in this imperfect human process.

[3] It is clear that the trend is getting worse, and that a large part of the delay is due also to requests for multiple revisions [6]. This is further interpreted as a cultural shift in the community away from coarse but interesting papers towards more polished papers [7]. It is also clear that reviews take a long time as papers have gotten longer, but that paper-length is not enough to account for this delay which primarily seems to come from the fact that it takes time before a paper is even read [8], [9].

[4] The surprise for economists is why the system has not already collapsed without such cash payments since on the surface, the reviewer would derive no private utility from reviewing. The answer has been to posit that the reviewer cares about the quality of the journal for some idiosyncratic reason [11].

Another class of systems where nodes simultaneously consume and provide services are peer-to-peer systems in networking. In peer-to-peer systems, the problem of freeloaders who can consume more resources than they "earn," has been long recognized and protocols like BitTorrent use explicit bartering (tit-for-tat) based incentives to enforce cooperative behavior [16]–[18] to some success within a single transfer. Reputation systems [19]–[21] have become another important topic of study in light of eBay's success, and it is natural to combine them with peer-to-peer systems to help create incentives for sharing even across different file transfers [22]–[24]. The tension between incentives and privacy has also been addressed a bit within the file-sharing community, but more in the context of ecash-like systems [25], [26] and Sybil attacks that rely on potentially cheap identities.

The existing literature, however, does not seem to address the potential tension between public reputations and the kind of anonymity that is desired in the context of peer-reviews. Since author identities and reputations for good work are decidedly expensive, there is no need to fear Sybil attacks. This makes the problem of scholarly peer review potentially easier. Although there are some indications that anonymous peer-review is not really necessary for quality purposes [27], scholarly tradition favors it greatly. In peer-to-peer filesharing systems, the true identity of a peer is already hidden behind an IP address, and protecting the IP address of a peer does not always arise as a social necessity[5].

Cash has the advantage of possibly motivating speedy reviews and doing so without being public (bank accounts are invisible). However, there are two major problems. The first is budgetary: for the most part, we simply cannot afford to pay enough to incentivize reviews.[6] The second issue is more subtle — by paying cash for reviews we run the real risk of destroying the "moral sentiments" of researchers. Samuel Bowles' recent survey [29] reviews how experimental economics strongly indicates that introducing monetary incentives often degrades higher ideals, sometimes irreversibly. With cash out of the running, it seems natural to study a public reputation based system. A diagnostic study of such non-financial incentives was recently conducted by the National Institute of Health (NIH) in the context of speeding up peer review of grants [30].

B. Our proposed system

Our proposed system for peer review is built upon a few basic hypotheses:

- Although it takes a nontrivial amount of time to perform a thorough review of a paper, a significant portion of the current delay in reviewing a paper is the time that elapses before the paper even gets read properly.
- Human beings are more likely to meet commitments and deadlines that are publicly proposed by themselves as compared to those that are imposed by others.
- In practice, reviewing papers can be roughly divided into two categories — "short-form reviews" that address the clarity, style, novelty, and interest-level of the paper, and "long-form reviews" that validate the correctness of the mathematical results in some detail. Short-form reviews take less time and are also more subjective. It is here that the experience and wisdom of the referee play a larger role. Long-form reviews are more objective in nature typically involving the correctness verification of technical contents.

At the heart of our proposal is a pair of centralized systems. The first tracks the score or "reputation" of any given person. The exact score will be made precise in the following section, but the idea is that it decreases with every paper submission and rises with every acceptable review. The score of each researcher is publicly available (in delayed or distorted versions), and it quantifies the extent to which the person is a good scholarly citizen who serves the community by performing reviews commensurate with the load imposed. The scores of students are clones of their respective faculty advisors' scores until they graduate, at which point they get their own identity.[7]

The second system is only semi-public. This allows a researcher to offer a commitment for a long or short form review. The commitment includes a starting date (at which point the paper is available to the reviewer and will presumably be read immediately) and an ending date (when the review is due). Assuming that researchers precommit to the review starting times based on their schedules, the length of this period is likely to be small, possibly around two to three weeks. The second system also maintains queues (*a la* Netflix) for each researcher consisting of papers that await his/her review. Papers are added to the queue by the editors in response to an accepted request for a review and can be removed at any time by either the reviewer or the editor. Papers in the queue are prioritized by the average[8] public scores of the authors[9] as read from the first system. At the beginning of

---

[5] If anonymous peering is desired, the traditional solution would be to use an anonymous routing strategy. Such systems have their own distinct problems of reputations [28] that are again distinct from the ones in peer review. In particular, for anonymity purposes, the system likes to have a lot of potentially "free-loading" traffic within which to hide the truly secret traffic.

[6] To get a quick sense of why this is, consider the IEEE Information Theory society. It currently has a structural surplus of about $100K per year and publishes roughly 1000 journal papers per year. Assuming that each needs only two reviews, that gives $50 per review. This is clearly not enough to motivate a behavior change. If authors were asked to pay a submission fee, this would almost surely be paid out of grants. A reviewing fee large enough to motivate behavior change — say matching consulting rates — would require at least $500 per review and would result in significant transfers of money from taxpayers to researchers that would not pass the "smell test." It would also raise problems with the educationally-useful practice of having graduate students and postdocs help with the reviews. A decent faculty member would then be compelled to share the reviewing money with the student. At this point, the faculty member would have to file extra tax forms and/or the student would have to deal with the taxes by treating this as self-employment. Would this run afoul of visa restrictions? All this just isn't practical.

[7] The purpose of this is clear — students will themselves partially reap the rewards of the papers they help review as students in the form of a higher score. It also creates another powerful incentive for faculty members to keep a high score — to avoid disadvantaging your own students.

[8] If a coauthor is the researcher's student, it counts only as one person.

[9] This provides another incentive to authors. If your score is low, you will end up being less attractive as a coauthor.

a review slot, the system sends the highest priority paper to the reviewer to read. There is also some indication to the editor of the expected time-of completion for the currently queued papers (subject of course to pre-emption by a higher priority paper) of any reviewer. This allows the editors to balance the delays across the different reviewers and out of self-interest, effectively offer people papers to review in proportion to their own desired number of papers to review.

Not every review request must be queued up within the second system. It is intended mostly for long-form reviews. Short-form reviews might very well be accommodated on an interrupt basis and we believe that the existence of such a system for getting long-form reviews might make editors more comfortable in asking for short-form reviews.[10] In the following sections, we consider a mathematical abstraction for such a system of long-form reviews and study the feasibility of the system by addressing two key questions: Does the system sufficiently incentivize the review process? Can the anonymity of reviewers be preserved under a public reputation system?

II. MATHEMATICAL MODEL

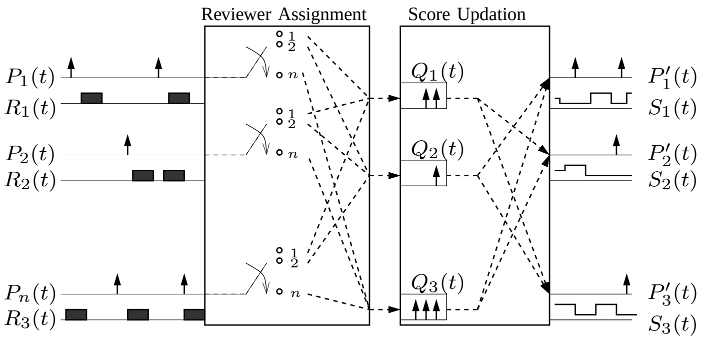


Fig. 1. Public Reputation based Review System. Queues correspond to the prioritized queue maintained for each reviewer in the system.

Researchers Consider a pool of n researchers working in an area. Researcher i submits papers for review at random times, which we model as a point process $P_i(t)$. The researcher also maintains a precommitted review slot schedule, modeled by a point process $R_i(t)$. The points of $R_i(t)$ represent the starting times of the review slots. The review duration is assumed to be a constant[11] D. The review reception times of papers submitted by researcher i is denoted by point process $P_i'(t)$. Note that the reviews may not arrive in the same order as the submitted papers.

Reputation/Score Based on the number of submitted and reviewed papers, each researcher has a time varying score that quantifies his/her service in the system. Specifically, we define a marked point process $\{\tau_{i,k}, M_k\}$ where $\tau_{i,k}$ is the union of points of the processes $P_i(t)$ and $R_i(t)$. The marker is an indicator if the point corresponds to a submitted or reviewed paper:

$$M_k = \begin{cases} 1 & \tau_{i,k} \text{ belongs to } R_i(t) \\ -1 & \text{o.w.} \end{cases} .$$

The score $S_i^L(t)$ of researcher i is defined as:

$$\begin{aligned} S_i^L(t) &= \alpha S_i(\tau_{i,m}) + \frac{1}{L}\sum_{j=m+1}^{m+L} M_j, \\ & \qquad \text{where } \tau_{i,m+L} \le t < \tau_{i,m+L+1}. \end{aligned} \tag{1}$$

The coefficient $\alpha \in [0,1]$ is a discount factor for the researchers' past activities, and L is a strictly positive integer that denotes the length (in ticks of the point process) of the researcher's current activity. If $\alpha = 0$, then $S_i(t)$ measures the (normalized) difference between number of submitted and reviewed papers neglecting the researchers activities prior to time $\tau_{i,m_i(t)}$.

Editor When a paper is submitted by researcher i for review, the editor finds K researchers to review the paper. The submitted papers are classified into a finite set of categories $\mathcal{C}$, based on the author[12], area of research and keywords. Given the paper's category and using the scores and earliest available review times of all researchers, the editor sends requests to other researchers until K affirmative responses are received[13]. For every submitted paper, the action of requests sent by the editor and the corresponding researchers' decisions to agree or disagree are combined into the single probability mass function $\{p(\mathbf{r}) : \mathbf{r} \subset \{1, \cdots, n\}, |\mathbf{r}| = K\}$, where $p(\mathbf{r})$ is the probability that the researchers in $\mathbf{r}$ have agreed to review the paper. In general, the probability mass function $\{p(\mathbf{r})\}$ depends on the category of the paper, the scores of researchers at time of submission, and the next available review slots of researchers. In the subsequent analysis, we provide certain conditions under which a review assignment can be termed as "fair" to the pool of researchers.

The very purpose of quantifying the service of a researcher is to incentivize the review process. If two papers are assigned to a particular researcher for review, the system ensures that the paper submitted by a researcher with higher priority is always assigned to an earlier slot than the one submitted by a lower-priority researcher.

III. INCENTIVIZING THROUGH PUBLIC REPUTATION: STABILITY AND DELAY

A. Homogenous Poisson Researchers

In order to gain insights about the functioning of such a reputation based system, we analyze the special case of a *homogenous researcher pool*, where all researchers have perfectly aligned interests, and every submitted paper is a

---

[10] Furthermore, useful unsolicited reviews of the preprints that appear on arXiv.org can also be given credit by the associate editors. This has the benefit of giving people some amount of proactive control over their scores without having to wait to be asked for a review.

[11] Constant review times are used merely for ease of presentation; our results can be extended to any delay distribution with bounded support.

[12] For the purpose of avoiding conflicts-of-interest and self-review, the author's identity is required to determine the appropriate reviewers.

[13] In reality, once a system like the one proposed is available, it might make sense to have some redundancy in the system by asking *more* than K reviewers and then removing the paper from their queues once enough reviews are received. For simplicity, we do not consider this case here.

single author paper in an identical area of study. In this case, $\mathcal{C} = \{1, \cdots, n\}$, and every paper submitted by researcher i belongs to category i. In general, the submission rates of researchers would be in an uncountable subspace of the positive reals. However, for analytical purposes, we consider a finite set of possible submission rates, and we divide researchers into a finite number of groups, such that all researchers in a group have identical paper submission rates. Let G be the total number of groups, and all researchers in group g submit papers according to independent Poisson processes of rate λ_g. Let $\{\mathcal{R}_g \subset \{1, \cdots, n\} : g = 1, \cdots, G\}$ denote the partition of researchers into the corresponding groups. We model the prespecified review slot schedule $R_i(t)$ of researcher i to be an independent Poisson processes of rate μ_i.

B. *Publication Stability under Priority Assignments*

At any given time t, the papers submitted by researchers to the editor that have not yet been reviewed, can be treated as a set of queues. Let $Q_i(t)$ denote the length of the queue containing papers submitted by researcher i, but not yet reviewed.

Definition 1: We define a researcher i to have **publication stability** if and only if the queue $Q_i(t)$ is stable.

The score, as defined in (1), is highly time-varying for any finite T, and since the review assignment is a function of the score, the system of queues would exist in a perpetually transient mode. To facilitate mathematical analysis of stability, we consider the steady-state score defined as:

$$S_i(t) = \lim_{\alpha \to 0, L \to \infty} S_i^L(t).$$

We assume an infinitely backlogged system, or in other words, every researcher always has a paper to review. For the pool of homogenous Poisson researchers, the steady state score for would then be given by the difference between their review and submission rates, normalized by the rate of the joint process: $S_i(t) = \frac{\mu_i - \lambda_i}{\mu_i + \lambda_i}$. Note that this would be the score had the authors been awarded points for review slots rather than completed reviews. Since $S_i(t)$ thus defined is a one-one function of the ratio $\frac{\mu_i}{\lambda_i}$ irrespective of time t, for the reminder of the stability analysis, we shall use $S'_i = \frac{\mu_i}{\lambda_i}$ to denote the score of researcher i.

Fairness in Review Assignment Since the choice of agreeing to review a paper is a researcher's prerogative, we consider a probabilistic model where reviewers are assigned independent of the next available slot times. In the special case of homogenous researchers with steady state scores, we consider the class of review assignment functions of the form $f_E : \mathbb{R}^n \times \mathbb{R}^n \times \mathcal{C} \to \mathcal{P}^{n,K}$, where $\mathcal{P}^{n,K}$ is the simplex of probability mass functions over cardinality K subsets of reviewers. If the list of scores $\mathbf{S} = \{s_1, \cdots, s_n\}$, the list of review slot rates $\mathcal{M} = \{\mu_1, \cdots, \mu_n\}$, then $f_E(\mathbf{S}, \mathcal{M}, i) = \{p(\mathbf{r})\}$ is the probability that the paper of category i is assigned to researchers in $\mathbf{r}$. The review assignment function should depend on the rates and scores in a manner that would ensure "fairness" in distribution of papers – as long as a researcher performs reviews commensurate to the submission rate, then his/her queue should be stable.

Specifically, we define a review assignment function f_E to be a *fair review assignment* if all researchers with scores greater than or equal to K have publication stability.

Theorem 1: In a homogenous Poisson pool of n researchers, let $S'_1 > S'_2 > \cdots S'_n$. Define $\mathcal{F}_E^k$ as the set of all review assignment functions f_E that satisfy the following criteria. Let $f_E(i, \mathbf{S}, \mathcal{M})\} = \{p_i(\mathbf{r})\}$ and $\mathcal{R}_g^k = \mathcal{R}_g \bigcup \{1, \cdots, k\}$.
1. For every $i \leq k$,

$$\sum_{\mathbf{r} \subseteq \{1, \cdots, k\}} p_i(\mathbf{r}) = 1.$$

2. For every $i \leq k$, $i \in g$

$$\sum_{\mathbf{r} \subseteq \{1 \cdots k\} : j \in \mathbf{r}} p_i(\mathbf{r}) = \frac{\lambda_g |\mathcal{R}_g^k|}{\sum_{g'=1}^{G} |\mathcal{R}_{g'}^k| \lambda_{g'}} \frac{\mu_j}{\sum_{l \leq k, l \neq i, l \in \mathcal{R}_g^k} \mu_l}.$$

3. For every $i > k$

$$\sum_{\mathbf{r} \subseteq \{1 \cdots k\} : j \in \mathbf{r}} p_i(\mathbf{r}) \leq \frac{\lambda_g |\mathcal{R}_g|}{\sum_{g'=1}^{G} |\mathcal{R}_g| \lambda_g} \frac{\mu_j}{\sum_{l \neq i} \mu_l}.$$

If $m = \arg\max\{i : S'_i \geq K\} > 1.$ and $|\mathcal{R}_g^m| \geq 2$, then any $f_E \in \mathcal{F}_E^m$ is a fair review assignment.

Proof: Refer to the Appendix. ☐

The above theorem states that there is a class of review assignments that guarantee fairness to researchers. The criteria that define the class of review assignment can be explained intuitively. First, the papers submitted by researchers who review commensurately with their submission rate, are only reviewed by those with a substantial review rate. Second, the group of a reviewer is first chosen with a probability proportional to the net arrival rate in that group. Within the group, the paper is assigned to a reviewer with a probability proportional to his/her standing in the group.

The strategy of assigning papers of the safe ($S'_i \geq K$) researchers within their pool is a conservative strategy that is sufficient for fairness. In general, since the review rates of some researchers would be higher than K, this pool of stable researchers can be expanded to include some lucky researchers whose scores are barely enough to share the demands of the high-scoring researchers, and can stand to benefit from the altruism of those researchers. Using the same class of review assignments from Theorem 1, the following theorem characterizes the size of this expanded stable researcher pool, and also provides the condition for instability.

Theorem 2: In a homogenous Poisson pool of n researchers, let $S'_1 > S'_2 > \cdots S'_n$ and let the number of groups $G = 1$. Under any review assignment in $\mathcal{F}_E^{M^*}$ where

$$M^* = \arg\max\{m : \sum_{i=1}^{m-1} \frac{1}{\sum_{j=1}^{m} S_j - S_i} \leq K\}, \tag{2}$$

all researchers in $\{1, \cdots, M^*\}$ have publication stability. A researcher i would not have publication stability if:

$$i \geq U^* \overset{\triangle}{=} \arg\min\{k : S_k \leq \frac{K}{1 - \sum_{i=1}^{M^*} \frac{K}{(\sum_{i=1}^{M^*} S_j) - S_i}} - \sum_{i \neq k} S_i\}.$$

Proof: Refer to the Appendix. □

Note that due to the definition of $\mathcal{F}_E^{M^*}$, the stable pool of researchers $\{1, \cdots, M^*\}$ in Theorem 2 are guaranteed publication stability irrespective of the scores of researchers $\{M^* + 1, M^* + 2, \cdots, n\}$. This *safe pool* contains some researchers who review fewer than K papers per submitted paper and yet have stability. Theorem 2 effectively divides the pool of researchers into four categories. The highest category of researchers are the *safe researchers* whose score exceeds K. As long as this pool is large enough, these researchers are guaranteed stability irrespective of what the actual scores of the researchers are. The next category are the *lucky researchers* who barely meet the criteria to enter the safe pool although their scores do not exceed K. These researchers are vulnerable to be removed from the safe pool if the scores of higher researchers decrease toward the safe threshold of K. Since the threshold for instability may not always be equal to $M^* + 1$, some researchers who do not have a sufficient score to enter the safe pool, might still be stable if there are enough residual slots in the system to guarantee their stability. These researchers belong to the category of *freeloaders* who are vulnerable to become unstable if the score of any researcher reduces. The last category is that of *unstable researchers* who face an unbounded delay in receiving reviews.

C. Why Increase Score: Delay Reduction

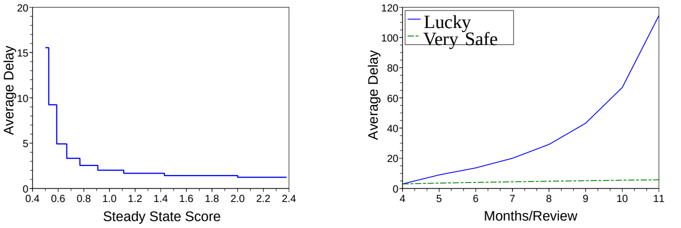
Although the minimum score required by a homogenous pool of researchers for guaranteed stability is K, one of the key incentives for increasing the score beyond the minimum is reducing the delay in receiving reviews. As the score of a researcher increases, his/her submitted papers are given higher priority at every reviewer's queue thereby reducing the overall delay in review reception. The following theorem characterizes mathematically the delays faced by researchers in the safe pool as a function of all their scores.

Theorem 3: In a stable pool of researchers $\{1, \cdots, M^*\}$, with scores $S'_1 > S'_2 \cdots > S'_{M^*}$, the average delay incurred by researcher $k \le M^*$ (when $K = 1$) is given by:

$$D_k = \frac{1}{\sum_{i\neq k, i\le m} S'_i\lambda}\sum_{i\neq k, i\le M^*}\left(\frac{1}{1-\sum_{j=1}^{k-1}\frac{1}{\sum_{l=1}^{k} S'_l - S'_j}}\right) \times \left(\frac{1}{1-\sum_{j=1}^{k-1}\frac{1}{\sum_{l=1}^{k} S'_l - S'_j} - \frac{1}{\sum_{l\neq k, l\le m} S'_l}}\right). \qquad (3)$$

Proof: Since the arrival and service times are exponentially distributed, the delay is a straightforward application of standard results in prioritized queuing systems, where an $M/M/1$ queue serves two arrival processes with different priorities. □

By scheduling review slots at a higher rate, any researcher can increase his/her score beyond the prevalent high score to be guaranteed highest priority and as a result, obtain the minimum possible delay in the system. This is evident from Figure 2.a, where the delay of a researcher is tracked as his/her score increases in a fixed pool of researchers. To further understand the benefit of higher scores when a group of researchers increase or decrease their scores, consider the application of Theorem 3 to the following example. Consider a stable pool of M researchers who submit papers at the rate of 1 paper every six months, and each paper is to be reviewed by $K = 1$ reviewer. Let $\frac{M}{2}$ researchers precommit to review slots at the rate of once every 4 months, while the other half commit at a rate μ less than once every 4 months. Then Figure 2.b plots the delays of the high priority (very safe) and low priority researchers as $\frac{1}{\mu}$ increases to the fair six-month threshold and beyond.



(a) (b)

Fig. 2. a) Delay versus score in a safe pool: $\lambda = 1$ review/6 months, μ ranges from 1 review/3 months to 1 review/20 months. b) Delay of group of researchers as their scores jointly decrease.

IV. ANONYMITY IN A PUBLIC REPUTATION SYSTEM

The proposed public reputation system provides additional information to authors about the activity of researchers at different points in time. This additional information obtained through submission and review reception times can be used to ascertain, or at the least narrow down, the set of possible reviewers for any particular paper. For example, if the proposed system updates the scores of reviewers (which are available in the public domain) instantaneously upon reception of a completed review, the identities fo all reviewers can be determined perfectly. Theorefore, unless the scores of reviewers are "distorted", no anonymity is achievable in the system. In this work, we study the achievable anonymity in a system where the scores of reviewers are allowed to be updated after a bounded delay.

Author as Eavesdropper Every author observes the processes $\{P_i(n), P'_i(n), S_i(n)\}$ which are time-discretized versions[14] of the processes $\{P_i(t)\}, \{P'_i(t)\}$ and $\{S_i(t)\}$ respectively from Figure 1. In other words, $X_i(n) = \int_{nT-T}^{nT} X_i(t)\mu(dt)$. We assume that an author who is serious about determining the identity of a reviewer would monitor these quantities for the entire duration of operation of the system. The authors know the probability distribution of assigning reviewers, and the (possibly random) strategy used in updating the scores of reviewers, but are unaware of the realization of the randomnesses involved.

---

[14] From a practical perspective, when arXiv entries, websites or journal footnotes are the sources of information, these quantities are indeed observable only in slots.

Score Updation: For a given delay constraint N, consider a deterministic strategy where scores of reviewers are updated periodically every N time slots. The author, therefore, is aware of the total number of reviews performed by each researcher within (periodic) $N-$slot windows.

Anonymity A key source of information to the authors is the order in which completed reviews are received. To understand this idea, consider a simple scenario where all researchers have identical scores which are high enough that every submitted paper gets reviewed in negligible time, and the order of reception of reviews is identical to the order of submission. Then, as the delay N in score updation increases, the number of reviews performed by each researcher would be nearly the same, thereby providing maximum anonymity. In general, the order of review reception does provide information about reviewer identities. However, as will be demonstrated in the subsequent discussion and simulations, this information becomes negligible as the updation delay increases.

Consider the joint paper submission process $P(n) = \bigcup P_i(n)$. For the j^{th} paper in $P(n)$, let $q_j(\mathbf{r})$ be the aposteriori probability that paper j was reviewed by the subset of researchers $\mathbf{r}$. The aposteriori probability is computed based on the complete observation of the author. Let Γ_j be the entropy:

$$\Gamma_j = -\sum_{r \subset \{1 \cdots n\}, |r| = K} q_j(\mathbf{r}) \log q_j(\mathbf{r}). \tag{4}$$

We define the anonymity $A(N)$ provided by the system is:

$$A(N) = \liminf_{J \to \infty} \frac{\sum_{j=1}^{J} \Gamma_j}{J \log \binom{n}{K}}. \tag{5}$$

The normalization in (5) ensures that the anonymity lies in $[0,1]$. $A(N) = 0$ implies that all reviewers are perfectly identified by every author, while $A(N) = 1$ implies that for every paper, the set of reviewers are equally likely to be any $K-$length subset of researchers.

The observations of the authors can be divided into independent cycles in time, where each *cycle* of observation begins when the first paper arrives into a system of empty queues (after an idle period of N slots), and the cycle ends when all queues are empty again for a period of N slots. For Poisson processes, the arrival and departure processes within different cycles are iid and it suffices to consider the observation within a generic cycle. Our goal is to demonstrate the efficacy of the simple delayed updation strategy, and for that purpose we focus on the scenario when each paper is assigned one reviewer ($K = 1$). It is intuitive that if $K > 1$, the reviewers can only have higher anonymity.

Let n_p be the number of papers that arrived in a cycle, and let $\mathbf{T}^p = \{t_1^p, \cdots, t_{n_p}^p\}$ be the arrival slots of the papers (wlog, $t_1 = 0$) within the cycle. Let $\mathbf{C} = \{c_1, \cdots, c_{n_p}\}$ denote the categories of the papers that arrived during the cycle. Let $\mathbf{T}^r = \{t_1^r, \cdots, t_{n_p}^r\}$ be the review reception slots of the papers within the cycle. We know that the updation slots are periodic, although an updation slot may not coincide with the start of a cycle. At each updation slot, the author observes an $n-$length vector containing the present scores of the n researchers. Let $\mathbf{U} = \{u_1, \cdots, u_{n_u}\}$ be the set of update vectors observed during the cycle. Therefore, the total observation of the author during the cycle is $\mathcal{Y} = \{\mathbf{T}^p, \mathbf{T}^r, \mathbf{u}, \mathbf{c}\}$, based on which the aposteriori probability of a paper being assigned to a reviewer can be computed as follows.

Let $\mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R})$ be the likelihood that review reception times equal $\mathbf{T}^r$ given the arrival times of papers $\mathbf{T}^p$ and the reviewer assignment is $\mathbf{R} = \{r_1, r_2, \cdots, r_n\}$ (r_i denotes the identity of the reviewer for paper i). Let $w(i, \mathbf{R}) = \sup\{j : j < i, r_j = r_i\}$. Then, $\mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R}) =$

$$\begin{cases} \prod_{i=1}^{n} \tilde{g}\left(t_i^r - \max(t_i^p, t_{w(i,\mathbf{R})}^r)\right) & t_i^r \geq \max(t_i^p, t_{w(i,\mathbf{R})}^r) \forall i \\ 0 & \text{o.w.} \end{cases}$$

where $\tilde{g}$ is the discrete-time approximation of the distribution of "inter-review" times:

$$\tilde{g}(k) = \begin{cases} \lambda e^{-\mu(k - \lceil \frac{D}{T} \rceil T}(1 - e^{-\mu T}) & k - \lceil \frac{D}{T} \rceil \geq 0, \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

Note that every realization of the pair $\mathbf{T}^r, \mathbf{R}$ would correspond to a unique sequence of updates $\mathbf{U}$. Therefore,

$$\mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R}, \mathbf{U}) = \begin{cases} \mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R}) & \mathbf{T}^r, \mathbf{R}, \mathbf{U} \text{ consistent} \\ 0 & \text{otherwise} \end{cases}$$

Using the above equation, the aposteriori probability that paper i was assigned to reviewer j is given by

$$\begin{aligned} q_i(j) &= \mathcal{L}(r_i = j | \mathbf{T}^r, \mathbf{T}^p, \mathbf{U}, \mathbf{C}) \\ &= \frac{\sum_{\mathbf{R}: r_i = j} \left(\mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R}, \mathbf{U}) \left(\prod_i p_i(r_i)\right)\right)}{\sum_{\mathbf{R}} \left(\mathcal{L}(\mathbf{T}^r | \mathbf{T}^p, \mathbf{R}, \mathbf{U}) \left(\prod_i p_i(r_i)\right)\right)}, \end{aligned}$$

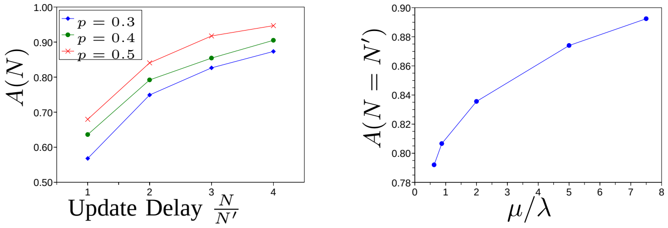
where $p_i(r_i)$ is the probability that a paper of category c_i is assigned to reviewer r_i (obtained from f_E, see Section III). The conditional entropy and the anonymity can then be computed using (4) and (5).

Using the derived expressions, we use simulations to demonstrate the gain in anonymity due to delayed updation. Specifically, consider a system where the total arrival rate of papers are according to Poisson process of rate λ. For ease of computation, we assume that each paper is reviewed by one of 2 reviewers, neither of whom are authors of any submitted paper.[15] Researchers commit to review slots according to Poisson processes of equal rate μ. The probability that any paper is assigned to researcher 1 for review is given by $p_i(1) = p \forall i$. From $M/M/1$ queue analysis, we know that the length of cycles grows exponentially with $\frac{1}{2\mu - \lambda}$. Hence, for computational purposes, we divide the cycle into time-periods of length N' slots, and mandate that the review reception times of papers that arrived within each N' slot time period fall within the same time-period (by suitably advancing review slots that cross over). Further, let the updation delay N be an integral multiple of N'. While this truncation only

---

[15] Note that in reality, the pool of reviewers would be much larger than 2, and this represents the bare minimum required for any positive anonymity to be achieved in a public reputation system.

approximates the original system, it is easy to see that as N', N increase, the difference in sample paths of the truncated and original systems becomes negligible.



(a) (b)

Fig. 3. a) Anonymity versus updation delay: $\lambda = 4/N', 2\mu = 5/N'$, $N' = 500$. b) Anonymity versus Score: $\lambda = 4, N' = 500$.

Figure 3.a plots the anonymity as a function of the ratio N/N'. This ratio represents the number of independent windows of observation between two successive score updates by the system. The anonymity increases with the increase in updation delay, and approaches the apriori entropy based purely on the review assignment probabilities. That the apriori entropy is an upper bound is immediately obvious, as the entropy conditioned on the observations is always less than the unconditional entropy (without having observed the score updates). Figure 3.b shows that anonymity increases as the review rates increase. The intuitive argument is that as review rates increase, there is less chance of completed reviews arriving out of order, which reduces the information available to the authors. This behaviour suggests that increasing anonymity is an additional incentive for researchers to have higher scores.

V. CONCLUDING REMARKS

This article has explored a way to incentivize good scholarly citizenship in the context of peer review — authors should review papers commensurate with the number of papers that they submit. To do this, a system of public reputations for authors has been proposed in combination with a peer-review system that gives a higher priority to authors who have reviewed relatively more papers. A crude analysis of the system shows that this indeed incentivizes reviewing to the extent that authors care about the reviewing delays that their papers experience. To maintain the anonymity of the reviewers, we argue that the scores should be distorted in some way. In this abstract, the process was distorted by sample-and-holding it at a slow enough rate. However, we have not analyzed the possible tension between this distortion and the author's desire for guaranteed low delay. This would probably require a transient analysis to complement the steady-state calculations here. We suspect that the distortion would cause authors to want to overprovision reviews slots to a small extent to give themselves a "safety margin."

APPENDIX: PROOFS

A. *Proof of Theorem* 1

Let $S'_1 > S'_2 > \cdots > S'_m > K > S'_{m+1} > \cdots S'_n$ be the scores of the n researchers. According to the condition, in $\{1, \cdots, m\}$, there are at least 2 researchers from each group g (groups with 0 researchers in the pool are precluded). Without loss of generality, let researcher 1 belong to group 1. Then, researcher 1 will have publication stability iff the virtual queues at every researcher in $\{2, \cdots, m\}$ containing only researcher 1s papers is stable under a review assignment in $\mathfrak{F}^m_{AE}$. Let $Q_{i,j}$ denote the queue at researcher j containing only researcher i's papers. Consider any $k \in \{2, \cdots, m\}$. The arrival process of researcher 1's papers into researcher k's queue is a Poisson process, specifically, a thinned version of the λ_1 process with the thinning coefficient given by $\sum_{S \subseteq \{2,\cdots,m\}: k \in S} p_1(S)$. Since researcher 1 has the highest priority, $Q_{i,k}$ will be stable iff the arrival rate of researcher 1 is less than the slot rate at researcher k. Let g_i denote the group of researcher i and let

$$I^k_{i,j} = \begin{cases} |\mathcal{R}^k_{g_j}| & g_j \neq g_i \\ |\mathcal{R}^k_{g_i}| - 1 & g_j = g_i \end{cases}$$

We divide this analysis into two cases depending on whether researcher k belongs to group $g_1 = 1$ or not. If researcher k belongs to group 1, then $Q_{1,k}$ is stable iff

$$\lambda_1 K \frac{\lambda_1 I^m_{i,k}}{\sum_{g=2}^G \lambda_g |\mathbb{R}_g| + \lambda_1 I^m_{1,k}} \frac{\mu_k}{\sum_{i \in I^m_{i,k}} \mu_i} \leq \mu_k$$

which is true if:

$$\sum_{i \in I^k_{i,k}} \mu_i \geq K\lambda_1 \frac{\lambda_1 I^k_{i,k}}{\sum_{g=2}^G \lambda_g |\mathcal{R}^m_g| + \lambda_1 (|\mathcal{R}_1|^m - 1)}.$$

Since $\mu_i \geq K\lambda_1$ for all $i \in \mathcal{R}^k_1$ and $|I_{1,k}| = |\mathcal{R}^m_1| - 1$, researcher 1 will have publication stability if

$$\frac{\lambda_1 I^m_{1,k}}{\sum_{g=2}^G \lambda_g |\mathcal{R}^m_g| + \lambda_1 (|\mathcal{R}^m_{g_1}| - 1)} \leq 1,$$

which is true. If researcher k belongs to group $l \neq 1$, then by replacing $\lambda_1(|\mathcal{R}_{g_1}| - 1)$ by $\lambda_l |\mathcal{R}_{g_l}|$ and $I^k_{1,k}$ by $I^k_{l,k}$ in the above argument, the proof follows.

Any researcher in $s \in \{2, \cdots, m\}$ will have publication stability if $\forall k \leq m$:

$$\frac{\lambda_{g_s} \lambda_{g_k} I^m_{s,k}}{\sum_g \lambda_g |\mathcal{R}^m_g|} \frac{\mu_k}{\sum_{i \in \mathcal{R}^m_{g_k}} \mu_i} \leq \frac{\mu_k}{K} - \sum_{i=1}^{s-1} \frac{\lambda_{g_s} \lambda_{g_k} I^m_{i,k}}{\sum_g \lambda_g |\mathcal{R}^m_g|} \frac{\mu_k}{\sum_{j \in \mathcal{R}^m_{g_k}} \mu_j}.$$

We know that

$$\forall i, k, \ \frac{\lambda_{g_k} I^m_{i,k}}{\sum_{j \in \mathcal{R}^m_{g_k}, j \neq i} \mu_j} \leq \frac{1}{K}.$$

Therefore, queue $Q_{s,k}$ would be stable if

$$\sum_{j=1}^{s} \frac{\lambda_{g_j}}{\sum_g \lambda_g |\mathcal{R}^m_g|} \leq 1. \qquad \square$$

B. *Proof of Theorem* 2

1. When $G = 1$, all researchers have identical paper submission rates λ. We know from the previous proof that if the number of researchers with score greater than K is at least 2, then the stable researcher pool is non-empty. It is easy to see that showing $M^* \geq i$ is equivalent to every pair of queues $Q_{i,j}, Q_{j,i}$ being stable; every researcher $j < i$ has a higher priority than i, and hence encounters a higher service rate than i at researchers $k < i$. Further, due to the prioritized and proportionate reviewer assignment, if $Q_{i,j}$ is stable then the sets of queues $\{Q_{k,j} : k \leq i\}$ and $\{Q_{i,k} : k \leq j\}$ are all stable. Therefore, to determine if i is an element of the stable pool, it is sufficient to consider the stability conditions of queues $Q_{i,i-1}, Q_{i-1,i}$. Queue $Q_{i,i-1}$ is stable iff:

$$\begin{aligned}\lambda K \frac{\mu_{i-1}}{\sum_{j=1}^{i-1} \mu_j} + \lambda K \sum_{s=1}^{i-2} \frac{\mu_{i-1}}{\sum_{j=1}^{i} \mu_j - \mu_s} &\leq \mu_{i-1} \\ \text{iff} \quad \sum_{s=1, s\neq i-1}^{i} \frac{1}{\sum_{j=1}^{i} S'_j - S'_s} &\leq \frac{1}{K}.\end{aligned}$$

Similarly $Q_{i-1,i}$ is stable iff $\sum_{s=1,s\neq i-1}^{i} \frac{1}{\sum_{j=1}^{i} S'_j - S'_s} \leq \frac{1}{K}$. Since the scores are strictly decreasing, it is easily shown that this condition subsumes condition (7). Furthermore, since the conditions are necessary and sufficient, the definition of M^* in the theorem represents the stable pool of researchers. □

2. According to the review assignment, the papers submitted by every researcher outside the stable researcher pool is proportionately assigned to all possible reviewers. Due to the proportionate assignment, and the fact that the size of the stable pool is determined by exhausting the slots of high-scoring researchers, the stability criterion that would be violated for researcher $k > M^*$ would correspond to a queue $Q_{k,j}$ where $j \leq M^*$. $Q_{k,j}$ would be unstable iff:

$$\lambda \frac{\mu_j}{\sum_{i=1}^{n} \mu_i - \mu_k} > \mu_j - \sum_{i=1}^{M^*} \frac{\mu_j}{\sum_{l=1}^{M^*} \mu_l - \mu_i} - \sum_{i=M^*+1}^{k-1} \frac{\mu_j}{\sum_{l\neq i}^{n} \mu_l}.$$

Since $\mu_i > \mu_{i+1} \forall i$, the above inequality holds if

$$\frac{k - M* -1}{\sum_{i=1}^{n} S'_i - S_k} > 1 - \frac{1}{1 - \sum_{l=1}^{M^*} \mu_l - \mu_i}.$$

Rearranging terms,the theorem is proved. □

ACKNOWLEDGMENTS

The research was partially supported by NSF grants CCF-0728872 and CCF-0729122.

REFERENCES

[1] C. Kerr, “View from the bridge,” *Time Magazine*, Nov. 1958.

[2] G. Hardin, “The tragedy of the commons,” *Science*, vol. 162, no. 3859, Dec. 1968.

[3] J. M. Campanario, “Peer review for journals as it stands today: part 1,” *Science Communication*, vol. 19, no. 3, pp. 181–211, Mar. 1998.

[4] ——, “Peer review for journals as it stands today: part 2,” *Science Communication*, vol. 19, no. 4, pp. 277–306, Jun. 1998.

[5] R. Smith, “Peer review: a flawed process at the heart of science and journals,” *Journal of the Royal Society of Medicine*, vol. 99, pp. 178–182, 2006.

[6] G. Ellison, “The slowdown of the economics publishing process,” *Journal of Political Economy*, vol. 110, no. 5, pp. 947–993, Oct. 2002.

[7] ——, “Evolving standards for academic publishing: a q-r theory,” *Journal of Political Economy*, vol. 110, no. 5, pp. 994–1034, Oct. 2002.

[8] A. Chong, “On the lags between submission and acceptance: are all referees created equal?” *Applied Economics Letters*, vol. 8, no. 6, pp. 423–425, 2001.

[9] D. S. Hamermesh, “Facts and myths about refereeing,” *The Journal of Economic Perspectives*, vol. 8, no. 1, pp. 153–163, 1994.

[10] O. H. Azar, “The review process in economics: is it too fast?” *Southern Economic Journal*, vol. 72, no. 2, pp. 482–491, 2005.

[11] M. Engers and J. Gans, “Why referees are not paid (enough),” *The American Economic Review*, vol. 88, no. 5, pp. 1341–1349, Dec. 1998.

[12] J. jen Chang and C. chong Lai, “Is it worthwhile to pay referees,” *Southern Economic Journal*, vol. 68, no. 2, pp. 457–463, Oct. 2001.

[13] Y. Riyanto and I. H. Yetkiner, “A market mechanism for scientific communication: a proposal,” *KYKLOS*, vol. 55, pp. 563–568, 2002.

[14] R. K. Goel, “A market mechanism for scientific communication: a comment,” *KYKLOS*, vol. 56, pp. 395–400, 2003.

[15] Y. Riyanto and I. H. Yetkiner, “A market mechanism for scientific communication: reply,” *KYKLOS*, vol. 56, pp. 401–404, 2003.

[16] B. Cohen, “Incentives build robustness in BitTorrent,” in *Proceedings of the 2nd IPTPS*, Berkeley, CA, Feb. 2003.

[17] D. Qiu and R. Srikant, “Modeling and performance analysis of bittorrent-like peer-to-peer networks,” *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, pp. 367–378, Oct. 2004. [Online]. Available: http://portal.acm.org/citation.cfm?id=1030194.1015508

[18] M. Li, J. Yu, and J. Wu, “Free-riding on bittorrent-like peer-to-peer file sharing systems: Modeling analysis and improvement,” vol. 19, no. 7, pp. 954–966, Jul. 2008.

[19] P. Resnick, R. Zeckhauser, E. Friedman, and K. Kuwabara, “Reputatation systems,” *Communications of the ACM*, vol. 43, no. 12, pp. 45–48, Dec. 2000.

[20] C. Dellarocas, “How often should reputation mechanisms update a trader’s reputation profile?” *Information Systems Research*, vol. 17, no. 3, pp. 271–285, Sep. 2006.

[21] M. Fan, Y. Tan, and A. B. Whinston, “Evaluation and design of online cooperative feedback mechanisms for reputation management,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 2, pp. 244–254, Feb. 2005.

[22] K. Ranganathan, M. Ripeanu, A. Sarin, and I. Foster, “‘to share or not to share’ an analysis of incentives to contribute in file sharing environments,” in *Proceedings of the Workshop on Economics of Peer-to-Peer Systems*, Jun. 2003.

[23] V. Vishnumurthy, S. Chandrakumar, and E. G. Sirer, “KARMA: A secure economic framework for P2P resource sharing,” in *Proceedings of the Workshop on the Economics of Peer-to-Peer Systems*, Berkeley, CA, Jun. 2003.

[24] H. T. Kung and C. H. Wu, “Differentiated admission for peer-to-peer systems: incentivizing peers to contribute their resources,” in *Proceedings of the Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, Jun. 2003.

[25] M. Belenkiy, M. Chase, C. C. Erway, J. Jannotti, A. Küpcü, A. Lysyanskaya, and E. Rachlin, “Making P2P accountable without losing privacy,” in *Proceedings of the 2007 ACM workshop on Privacy in electronic society*. New York, NY, USA: ACM, 2007, pp. 31–40.

[26] S. Marti and H. Garcia-Molina, “Identity crisis: anonymity vs reputation in P2P systems,” in *Proceedings of the Third International Conference on Peer-to-Peer Computing*, Sep. 2003, pp. 134–141.

[27] F. Godlee, C. R. Gale, and C. N. Martyn, “Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial,” *Journal of the American Medical Association*, vol. 280, no. 3, pp. 237–240, Jul. 1998.

[28] R. Dingledine, N. Mathewson, and P. Syverson, “Reputation in P2P anonymity systems,” in *Proceedings of Workshop on Economics of Peer-to-Peer Systems*, Berkeley, CA, Jun. 2003.

[29] S. Bowles, “Policies designed for self-interested citizens may undermine ‘the moral sentiments’: evidence from economic experiments,” *Science*, vol. 320, Jun. 2008.

[30] “NIH peer review report final draft,” Jul. 2008. [Online]. Available: http://enhancing-peer-review.nih.gov/meetings/NIHPeerReviewReportFINALDRAFT.pdf