An Automatically Refereed Scholarly Electronic Journal: Formal Specifications

Stefano Mizzaro and Paolo Zandegiacomo Riziò
Department of Mathematics and Computer Science
University of Udine
Via delle Scienze, 206 — Loc. Rizzi, I 33100 Udine, Italy
e-mail: mizzaro@dimi.uniud.it, zandegia@tagliamento.sci.uniud.it
www: http://www.dimi.uniud.it/~mizzaro

Keywords: Digital libraries, scholarly journals, electronic publishing, peer review, quality control

Edited by:

Received: **Revised:** **Accepted:**

Internet growth seems to amplify the critiques to peer review mechanism: many researchers maintain that Internet would allow a faster, more interactive, and more effective model of publishing. However, just removing peer review would lead to a lack of quality control in scholarly publications. We propose a new kind of electronic scholarly journal, in which the standard submission-review-publication process is replaced by a more democratic approach, based on judgments expressed by the readers. The new electronic scholarly journal is described in both intuitive and formal ways.

1 Introduction: scholarly journals and peer review

The communication mechanism that modern science still adopts nowadays arose in the 17th Century, with the publication of the first scientific journals reporting, in paper form, the ideas, discoveries, inventions, of researchers. Nowadays, since about 1930, the dissemination of scholarly information is based on *peer review*: the researcher that wants to disseminate her work writes a paper and submits it to a scholarly journal; the paper is not immediately published, but it is judged by some referees; if they judge it adequate, the paper is published.

The peer review mechanism ensures a reasonable *quality* of the published papers, and it is usually retained an adequate solution, though not the ideal one. Indeed, the peer review mechanism can be (and has been) criticized. Sometimes, the reviewing process takes too long, even one or two years, so that the published paper describes something old. Sometimes the reviewers do not do a good job, accepting a bad paper or not accepting a good one, that after two years cannot be resubmitted because too obsolete. Sometimes, referees introduce some bias in published papers: for instance, in medicine field, papers describing negative results seem more difficult to publish than paper describing positive ones. And one might go on.

Internet has changed, and is changing, this situation [1, 2, 7, 11]. A peer reviewed journal can be distributed by electronic means. The refereeing process too can take place completely electronically, drastically reducing time and money: see, *e.g.,* JHEP (http://jhep.sissa.it) or Earth Interactions [5] (http://EarthInteractions.org). Multimediality can lead to a more effective communication [5]. Of course, there are also some drawbacks of electronic journals (copyright problems, legal validity, accessibility, and so on), and they seem not to have a large impact by now [4], but the general feeling is that this is a temporary situation, and we just have to wait some years for overcoming these temporary problems.

Internet growth seems to amplify the critiques to peer review: many researchers maintain that Internet would allow a more fast, elastic, interactive, and effective model of publishing. Nadasdy [10] suggests substituting peer review with democracy: each submitted paper is immedi-

ately published and readers will judge it, selecting what they deem useful. Of course, the problem with this approach is that the readers may not be capable of correctly judging the paper: whereas the referees are chosen among the experts in the field, everybody can read and judge a paper published on Internet.

Nadasdy's proposal is not an abstract one. In a few years, this model of publishing might become a *de facto* standard, as witnessed by two examples, already existing today: the "do it yourself publishing" (authors publishing in a web site their ideas), and public repositories of scholarly papers (where authors can publish papers classified in some categories—see, *e.g.,* http://ArXiv.org). The threat that these (without-quality-control) mechanisms will replace the (with-quality-control) peer review journals is a real one.

A proposed solution is to replace peer review with *peer commentary*: readers will write in a public commentary their judgments on the read papers. It seems a viable solution, but Harnad, after some practical experience with this solution, says that "peer commentary is a superb supplement to peer review, but it is certainly no substitute for it" [3].

We propose a more sophisticate mechanism. We describe a new kind of electronic scholarly journal, with the aim of changing the submission-review-publication process, making it more automatic, keeping anyway at a high level the quality of scientific papers, and providing also a way of measuring in an automatic and objective way the quality of *researchers*, extending and improving the well known *impact factor* mechanism (http://www.isinet.com/hot/essays/7.html). We try to make a step further on the road suggested by the not refereed journals just mentioned, and to present a mechanism that avoids some of the previously described problems.

This paper, that extends and revises previous work [8, 9], is structured as follows. In Section 2, the mechanism is described in an intuitive way. In Section 3, the behavior of the whole system is formally defined by means of some formulæ. Section 4 discusses some open problems and future developments.

2 General description

The basic idea is the following. Imagine a scholarly journal in which each paper is immediately published after its submission, without a refereeing process. Each paper has a score, measuring its quality. This score is initially zero, and is later dynamically updated on the basis of the readers' judgments. A subscriber of the journal is an author or a reader (or both). Each subscriber has a score too, initially zero and later updated on the basis of the activity of the subscriber (if the subscriber is both an author and a reader, she has two different scores, one as an author and one as a reader). Therefore, the scores of subscribers are dynamic too, and change accordingly to subscribers' behavior: if an author with a low score publishes a very good paper (*i.e.,* a paper judged very positively by the readers), her score increases; if a reader expresses an inadequate judgment on a paper, her score decreases accordingly, and so on.

Every object with a score (author, reader, paper) has also a *steadiness* value, that indicates how much steady the score is: for instance, old papers will have a high steadiness; new readers (authors) will have a low steadiness. Steadiness affects the score update: a low (high) steadiness allows quicker (slower) changes of the corresponding score. A steadiness value increases as the corresponding score changes.

While time goes on, readers read the papers, judgments are expressed, and the corresponding scores and steadinesses vary consequently. The score of a paper can be used for deciding to read or not to read that paper; the scores of authors and readers are a measure of their research productivity, then they will try to do their best for keeping their score at a high level, hopefully leading to a virtuous circle (publishing good papers and giving correct judgments to the read papers). A steadiness value is an estimate of how stable (and, therefore, reliable) the corresponding score is.

For understanding the details of the automatically refereed journal proposed here, let's follow the events happening when a paper is read and judged by a reader. The following scores and steadinesses change:

– **Paper.** The paper score is updated: if the

judgment is lower (higher) than the actual paper score, the paper score decreases (increases). The score of the reader determines the weight of the judgment: judgments given by higher rated readers will be more important (will lead to higher changes) than judgments given by lower rated readers.

The steadiness of the paper increases, since the score of the paper is now computed on the basis of one more judgment, and is therefore statistically more reliable.

- **Author.** The author's score is updated: when the score of a paper written by an author decreases (increases), the score of the author decreases (increases). Thus, authors' scores are linked to the scores of their papers.

  The steadiness of the author increases, since the score of the author is now obtained with one more judgment and is therefore statistically more reliable.

- **Reader.** The reader's score is updated: if one reader's judgment about a document is "wrong" (too far from the average), the reader's score has to decrease. Then, the reader's score is updated depending on the goodness of her judgment (how much adequate her judgment is, or how much it agrees with the current score of the paper).

  The steadiness of the reader increases, since her score, computed on the basis of the goodness of her judgments, is obtained on the basis of one more judgment.

- **Previous readers.** The scores of the readers that previously read the same paper are updated: if a judgment causes a change in a paper score, all the goodnesses of the previously expressed judgments on that paper have to be re-estimated. Therefore, a judgment on a certain paper leads to an updating of the scores of all the previous readers of that paper.

  The steadinesses of the previous readers increase since the goodnesses of the readers, that lead to their scores, are obtained on the basis of one more judgment.

The updating of the scores of the previous readers deserve further explanation. After the paper

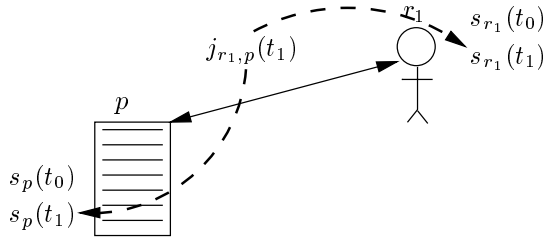

Figure 1: The updating of previous readers' scores: t_1.

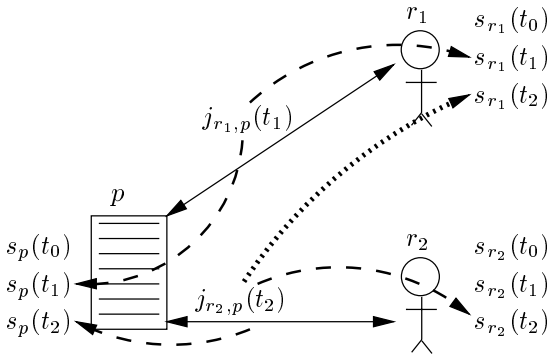

Figure 2: The updating of previous readers' scores: t_2.

score has changed, it is possible to revise the goodness of the old readers' judgments, and to update the old readers' score consequently: for instance, if an old reader r expressed a judgment j that was "bad" (distant from the paper score) at that time, but after that the paper score changes and becomes more similar to j, then the score of r (s_r) has to increase. Let us take into account a simple concrete example (see Figures 1, 2, and 3):

- At time t_0, we have a paper p with score $s_p(t_0)$, three readers r_1, r_2, and r_3 with their scores $s_{r_1}(t_0)$, $s_{r_2}(t_0)$, and $s_{r_3}(t_0)$.

- At the following time instant $t_1 > t_0$ (Figure 1), reader r_1 reads paper p expressing the judgment $j_{r_1,p}(t_1)$ (continuous double arrow line in figure). This causes the updating of the scores of p and r_1 (dashed line in figure): we obtain $s_p(t_1)$ and $s_{r_1}(t_1)$.

- At time $t_2 > t_1$ (Figure 2), reader r_2 reads p expressing $j_{r_2,p}(t_2)$. The scores of p and r_2 are updated consequently, leading to $s_p(t_2)$ and $s_{r_2}(t_2)$. But also the score of r_1 has

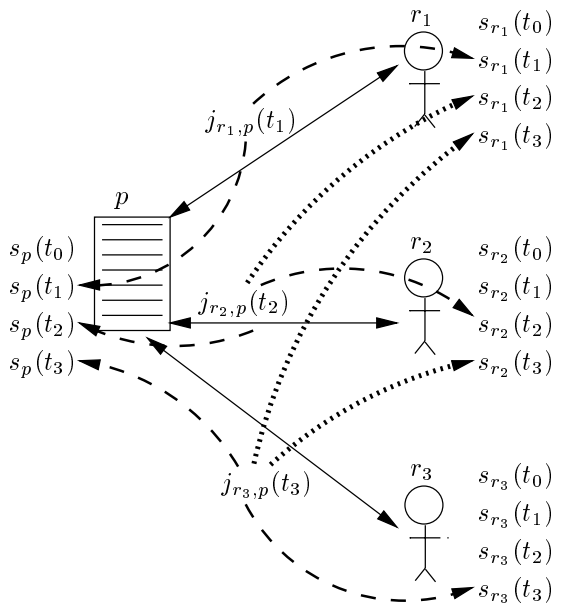

Figure 3: The updating of previous readers' scores: t_3.

to be updated (dotted line in figure), since the goodness estimated at time t_1 for $j_{r_2,p}(t_2)$ with respect to $s_p(t_1)$ has to be re-estimated now that the score of p is $s_p(t_2)$.

– At time $t_3 > t_2$ (Figure 3), r_3 reads p expressing $j_{r_3,p}(t_3)$. This changes the score of p ($s_p(t_3)$), the score of r_3 ($s_{r_3}(t_3)$), and the scores of the previous two readers ($s_{r_2}(t_3)$ and $s_{r_1}(t_3)$).

In other words, the goodness of a reader's judgment is an approximation of the *ideal goodness*, defined as the difference between the reader's judgment and the final score of the paper (*i.e.*, the score obtained when the last judgment on that paper has been expressed). Since the final score is obviously not available when the judgment is expressed, it has to be guessed (updating of the reader), but this guess is revised and refined as time goes on and tends to $= +\infty$ (updating of previous readers).

3 Invariant properties

In this section, we will present some formulæ, in order to formally specify how to compute the values of the scores and steadiness of paper, author, reader, and previous readers, depending on the expressed judgments.

Let's start with some notation. We will denote with:

– t and t_i the discrete time instants. We assume that t_{i+1} immediately follows t_i, and that between t_i and t_{i+1} only the explicitly specified events will happen.

– $s_r(t), s_p(t), s_a(t)$ the score of a reader, a paper, and an author, respectively, at time t. We will sometimes omit the time indication when this does not rise ambiguity. The values for $s_p(t)$ and $s_a(t)$ are in the range $[0, 1]$ (0 is the minimum and 1 the maximum), whereas the values for $s_a(t)$ are in the range $]0, 1]$. This difference will be explained in Section 3.3.

– $\sigma_r(t), \sigma_p(t), \sigma_a(t)$ the steadiness of a reader, a paper, and an author, respectively, at time t. All the steadiness values are in the range $[0, +\infty[$.

– $j_{r,p}(t)$ the judgment expressed at time t by reader r on paper p.

– $t_{r,p}$ the time instant of the judgment expressed by r on p (we are implicitly assuming that each reader can judge each paper only once). We will write $j_{r,p}$ instead of $j_{r,p}(t_{r,p})$.

3.1 Paper

Given a paper p, its score is the weighted mean of the judgments previously expressed by readers on p. The weight of each judgment is the score that the reader has when she expresses the judgment, to give more importance to the judgments given by better readers.

Definition 1 (Paper score s_p) Given a paper p, the set $R_p(t)$ of readers that expressed a judgment on p before time t, and the time instants of judgments expressions $t_{r,p}$, we have $\forall r \in R_p(t)$ the judgment $j_{r,p}$ expressed by r on paper p and the score $s_r(t_{r,p})$ of r at time $t_{r,p}$, *i.e.*, the score that the reader has when she expresses the judgment. We can then define the score of paper p at

time t as:

$$s_p(t) = \frac{\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p}}{\sum_{r \in R_p(t)} s_r(t_{r,p})}. \quad (1)$$

■

Remark 1 Consistently with Formula 1, the score of the paper before any judgment is expressed on it is zero.

Remark 2 The score $s_p(t)$ of a paper is modified only when a judgment on p is expressed.

Remark 3 In Formula 1, each judgment is weighted on the basis of the score that the reader has when she expresses the judgment ($s_r(t_{r,p})$). The alternative of using the score that the reader has "now", *i.e.*, when the mean is calculated ($s_r(t)$), seems less preferable since the reader's competence has probably changed in this period.

Let's now see how to measure the steadiness of a paper. The steadiness of a paper has to measure how stable its score is. To define it, a first attempt might be as the number of judgments expressed on that paper. However, it seems reasonable that a judgment expressed by a good reader should be more important, and give more steadiness to the paper, than a judgment expressed by a reader with a low score. Therefore, we define the steadiness of paper p at time t as the summation of the scores that readers have when they express their judgments on p.

Definition 2 (Paper steadiness σ_p) Given a paper p, the set $R_p(t)$ of readers that expressed a judgment on p before time t, and the time instants of judgment expression $t_{r,p}$, with $t_{r,p} \leq t$, the steadiness of p at time t is:

$$\sigma_p(t) = \sum_{r \in R_p(t)} s_r(t_{r,p}). \quad (2)$$

■

Remark 4 The steadiness value of a just published, and not yet judged, paper is zero.

Remark 5 The expression in the denominator in Formula 1 is exactly the steadiness of the paper. Therefore, we can rewrite Formula 1 as:

$$s_p(t) = \frac{\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p}}{\sigma_p(t)}. \quad (3)$$

3.2 Author

Given an author a, her score at time t can be defined in two equivalent ways:

- As the weighted mean of the scores of the papers previously published by a. The weight of each paper p is the steadiness of p, a value that sums up all the scores of the readers that expressed a judgment on p (see Formula 2).
- As the weighted mean of the judgments previously expressed by readers on the papers published by a. The weight of each judgment is, again, the score that the reader has when she expresses the judgment.

Let's define formally the first alternative, that uses the steadiness of a paper to weight the papers scores.

Definition 3 (Author score s_a) Given an author a and the set $P_a(t)$ of papers published by a before time t, we have $\forall p \in P_a(t)$ the score $s_p(t)$ of p at time t and the steadiness $\sigma_p(t)$ of p at time t. We can now define the score of author a at time t as:

$$s_a(t) = \frac{\sum_{p \in P_a(t)} \sigma_p(t) \cdot s_p(t)}{\sum_{p \in P_a(t)} \sigma_p(t)}. \quad (4)$$

■

Following the second alternative, we can define:

$$s_a(t) = \frac{\sum_{p \in P_a(t)} \left(\sum_{r \in R_p(t)} s_r(t_{r,p}) \cdot j_{r,p} \right)}{\sum_{p \in P_a(t)} \left(\sum_{r \in R_p(t)} s_r(t_{r,p}) \right)}, \quad (5)$$

where $P_a(t)$ is the set of papers published by a before time t, $R_p(t)$ is the set of readers that judged

paper p before t, $s_r(t_{r,p})$ is the score of r at time $t_{r,p}$, $t_{r,p} \leq t$ are time instants of judgment expression, and $j_{r,p}$ is the judgment expressed by r on paper p.

Remark 6 Using Formulæ 2 and 3 to rewrite the summations in parentheses in Formula 5, it is easy to see that the score of an author computed with Formulæ 4 and 5 are equivalent.

Remark 7 Consistently with Formulæ 4 and 5, the score of the author before any judgment is expressed on her papers is zero.

Remark 8 $s_a(t)$ is modified only when the score of one of the papers published by a changes, *i.e.*, when a judgment on a paper published by a is expressed.

Remark 9 As discussed in Remark 3, in Formula 5 each judgment is weighted on the basis of the score that the reader has when expresses the judgment ($s_r(t_{r,p})$).

The steadiness of an author has to measure how stable her score is. We can define it equivalently in two ways, as the summation of the steadiness of her papers and as the summation of the scores that the readers had when they expressed a judgment on a paper of the author. The first alternative leads to the following definition.

Definition 4 (Author steadiness σ_a) Given an author a, the set of papers $P_a(t)$ published by a, and the steadiness $\sigma_p(t)$ of each paper $p \in P_a(t)$ at time t, the steadiness of author a at time t is:

$$\sigma_a(t) = \sum_{p \in P_a(t)} \sigma_p(t). \quad (6)$$

■

Following the second alternative, we can define (with the usual notation):

$$\sigma_a(t) = \sum_{p \in P_a(t)} \sum_{r \in R_p(t)} s_r(t_{r,p}). \quad (7)$$

The equivalence of Formulæ 6 and 7 follows immediately from Formula 2.

Remark 10 Since the denominator in Formula 4 is exactly the steadiness of the author (Formula 6), we can rewrite Formula 4 as:

$$s_a(t) = \frac{\sum_{p \in P_a(t)} \sigma_p(t) \cdot s_p(t)}{\sigma_a(t)}. \quad (8)$$

3.3 Reader

First of all, we need to define a measure of how good a judgment is.

Definition 5 (Goodness) Given the score $s_p(t)$ of a paper p at time t and the judgment $j_{r,p}$ by reader r on p, the goodness, calculated at time t, of the judgment $j_{r,p}$ is defined as:

$$g_{j_{r,p}}(t) = 1 - |j_{r,p} - s_p(t)|. \quad (9)$$

■

Remark 11 The goodness depends on the time instant at which it is computed (t in Formula 9), since s_p can change also after the judgment expression (because of other judgments).

Given a reader r, her score is the weighted mean of the goodness of the judgments she has previously expressed. If we gave the same weight to all the goodnesses, we might define:

$$s_r(t) = \frac{\sum_{p \in P_r(t)} g_{j_{r,p}}(t)}{|P_r(t)|} \quad (10)$$

where $s_r(t)$ is the score of reader r at time t, $P_r(t)$ is the set of papers judged by r before t, $|P_r(t)|$ is the cardinality of $P_r(t)$, $j_{r,p}$ is the judgment expressed by r on paper p, and $g_{j_{r,p}}(t)$ is the goodness, calculated at time t, of the judgment $j_{r,p}$.

However, the weight of each goodness should be different on the basis of the steadiness of the paper being judged. For instance, a wrong judgment on a paper that has been previously judged by many readers should weight more than a wrong judgment on a paper whose score is calculated on the basis of very few judgments. The steadiness of the paper being judged should be taken into account, and the previous Formula 10 for $s_r(t)$ has to be rewritten as follows.

Definition 6 (Reader score s_r) Given a reader r and the set $P_r(t)$ of papers judged by r before time t, we have $\forall p \in P_r(t)$ the steadiness $\sigma_p(t)$ of p at time t, the judgment $j_{r,p}$ expressed by r on p, and the goodness $g_{j_{r,p}}(t)$ of the judgment $j_{r,p}$, calculated at time t. We can define the score of reader r at time t as:

$$s_r(t) = \frac{\sum_{p \in P_r(t)} \sigma_p(t) \cdot g_{j_{r,p}}(t)}{\sum_{p \in P_r(t)} \sigma_p(t)}. \quad (11)$$

■

Remark 12 The initial value of s_r should be, accordingly to Formula 11, zero. But this would lead to a division by zero in Formula 1 (if we compute the score of a paper that has not been judged yet). We therefore define this initial value as $s_r = \epsilon$, where ϵ is a small value that will be neglected after some judgments are expressed.

Remark 13 The score of a reader r changes when:

- r judges a paper. This adds one element to the set $P_r(t)$ and one new addendum to the summations in Formula 11.
- A reader $r' \neq r$ judges a paper that has previously been judged by r. This causes the score of the paper to change, and this leads to changing the goodness of the judgment previously expressed by r.

Remark 14 All the goodnesses are calculated at time t, to have the better estimate that is possible. We are implicitly assuming that the estimate improves with the number of judgments.

Remark 15 For the same reason, the steadinesses are calculated at time t too: in Formula 11, we have $\sigma_p(t)$, not $\sigma_p(t_{r,p})$.

The steadiness of a reader has to measure how stable the score of a reader is. A good measure might be the number of judgments expressed by a reader. However, since a judgment expressed on a paper with a high steadiness value should give more steadiness to the reader, we define the steadiness of a reader as the summation of the steadinesses of the papers judged by her.

Definition 7 (Reader steadiness σ_r) If $P_r(t)$ is the set of papers judged by a reader r before time t, the steadiness of r at time t is

$$\sigma_r(t) = \sum_{p \in P_r(t)} \sigma_p(t). \quad (12)$$

■

Remark 16 The denominator in Formula 11 is the steadiness of the reader (right hand side in Formula 12). Therefore, Formula 11 can be rewritten as:

$$s_r(t) = \frac{\sum_{p \in P_r(t)} \sigma_p(t) \cdot g_{j_{r,p}}(t)}{\sigma_r(t)}. \quad (13)$$

Remark 17 Although the steadiness of a reader is not used in the other formulæ, it is however useful to judge the reliability of a reader.

4 Conclusions and future developments

We have described, both intuitively (Section 2) and formally (Section 3), an electronic scholarly journal in which the standard peer review process is replaced by a more democratic approach based on judgments expressed by the readers.

The formulæ in Section 3 are different from what presented in [8, 9]: the formulæ presented here define some invariant properties of the system, that must hold at each time instant, whereas the formulæ in [8, 9] define how to update the values of score and steadiness of papers, authors, and readers as time goes on. The invariant properties are simpler than the updating formulæ, and thus it is straightforward to understand and justify them from an intuitive point of view; the correctness of the updating formulæ is more difficult to grasp at an intuitive level, but the equivalence of the two approaches can be formally proved. Moreover, because of some improvements and simplifications, the formulæ presented here are not fully equivalent to those in [8, 9].

Generally speaking, this proposal can be seen as an improvement of the dissemination of scholarly information through on line journals. More specifically, it can be seen as an improvement of the well known impact factor mechanism, of

the democratic journal proposed in [10], and of collaborative information retrieval and filtering [6], since it allows distinguishing among "good" and "bad" collaborators. Two other proposals that are similar, but less formalized, are those by Stodolsky [12] and by Varian [13].

This proposal is not free from problems. In general, one may wonder if democracy is a good approach to knowledge dissemination. Of course, it is difficult to have an objective opinion on that: it could be appropriate, or appropriate in some fields only, or not appropriate at all. We believe that only by further studies and experiments we can find an objective answer. However, it has to be emphasized that the mechanism proposed here is different from democracy, since different readers will have different importance. More specific problems are briefly discussed in the following.

A problem is *lazy readers*: a reader can simply confirm the previously expressed judgments, giving to each read paper a score equal to its actual score, with the aim of obtaining a high goodness. Two solutions seem suited here: give higher scores to fast readers (those that first read the papers), and do not show the paper score for a period after its publication (for instance, until when its steadiness reaches a certain value). Moreover, the strategy adopted by such a lazy reader might be a bad one, and lead to a low score for her, if the judgments by the following readers change the score of the paper.

Another problem might be the *lobbies*, *i.e.,* people that agree in mutually giving high scores. This might not be a problem at all, if the whole system can be modified to behave, by choosing appropriate formulæ and constants, in a way that discourages the lobbies; if this is not the case, the solution would probably be to implement some software able to detect such situations.

Of course there are technical difficulties too, *e.g.,* the identification of subscribers, or the huge amount of storage needed for recording the papers, the subscribers' data, and the history of expressed judgments. But these can be surely handled by database and cryptography technologies.

It is also easy to see some mandatory *improvements*:

- To deal with papers with more than one author. That should be easy: the judgments on the paper cause a modification of all the authors' scores, weighted on the basis of the importance of the contribution of each author to the paper.
- To have more scores, both for subscribers (authors and readers) and papers: comprehensibility, technical soundness, originality, and so on. In this way, a more detailed evaluation is available. If just one single number is needed, a weighted mean of all the scores of a subscriber (or a paper) can be used.
- To have more than one journal, with different acceptance thresholds: an author must have a score larger than the threshold for publishing. Younger researcher will subscribe to lower rated journals, and "first class" journals will accept only well established researchers. The mechanism presented in this paper can also be used as a complement of, instead of a replacement to, peer review: the initial score of a paper can be given by a standard peer review, thus allowing an author with a low score to submit her paper to an higher rated journal.
- To allow the subscribers know why their score is decreasing, *i.e.,* which "wrong" judgment, or paper, causes that, and eventually let them revise their judgment or withdraw their paper.
- To introduce some sort of *rent* function, for decreasing the score of subscribers that are inactive for long periods of time.
- To allow the readers to express, besides the numerical judgments also a free text *commentary* on the paper. The commentary can then be considered as a paper itself, and judged by other readers, but it is linked to the paper it comments, and the score of the commentary can affect the score of the paper.

Finally, we sketch how we intend to proceed with this research. A software simulator of the electronic journal proposed here is currently under implementation. We also plan to use some mathematical models and techniques for formally studying the behavior of the whole system and for studying other similar approaches. More in detail, some typical cases could be simulated, *e.g.,* what happens to a reader's score when she expresses a

judgment that is correct but far from the average; it would be possible to understand if and how the order of the judgments expressed by the readers affect their final score, or if the initial conditions—*i.e.*, the number of initial readers, their score, and so on—are a critical factor for having a stable system; we might use a bet-like approach, in which each reader has some money and bets on some papers; and so on.

These theoretical and experimental activities will allow to verify that the behavior of the system is correct and consistent and to choose in a more reliable way among the possible formulæ and parameters. After that, the software for the complete system will be implemented, tested, and evaluated. An ideal environment for these experiments is a repository of preprints, like ArXiv (http://ArXiv.org/). We plan to execute some laboratory experiments (with simulated papers, authors, and readers) and some real life experiments (involving real users), mostly within the TIPS project (http://tips.sissa.it), recently funded in the 5th Framework Program of the EU.

These theoretical and experimental activities will also allow to determine the relations between the values of some parameters and the real situation: indeed, it is likely that the values of parameters depend on figures like the number of subscribers, the rate of papers publishing, the rate of judgments expression, and so on. These dependencies must be singled out and, in case the above presented formulæ turned out to be inadequate, new ones need to be proposed and evaluated. Again, ArXiv seems the ideal environment for this issue.

Acknowledgements

We would like to thank Paolo Coppola and Luca Di Gaspero for having read a draft of this paper and provided many useful comments.

References

[1] C. W. Bailey, Jr. Scholarly electronic publishing bibliography—Version 32: 8/1/00. Web document, visited 3/9/00. http://info.lib.uh.edu/sepb/sepb.html.

[2] C. W. Bailey, Jr. Scholarly electronic publishing on the internet, the NREN, and the NII: Charting possible futures. *Serials Review*, 20(3):7–16, 1994. http://info.lib.uh.edu/cwb/schpub.htm.

[3] S. Harnad. Learned inquiry and the net: The role of peer review, peer commentary and copyright. *Antiquity*, 71:1042–1048, 1997. ftp://ftp.princeton.edu/pub/harnad/Harnad/HTML/harnad97.antiquity.html.

[4] S. P. Harter. Scholarly communication and electronic journals: An impact study. *Journal of the American Society for Information Science*, 49(6):507–516, May 1998.

[5] J. Holoviak and K. L. Seitter. Earth Interactions—Trascending the limitations of the printed page. *JEP the Journal of Electronic Publishing*, 3(1), 1997. ISSN 1080-2711; http://www.press.umich.edu/jep/03-01/EI.html.

[6] M. Karamuftuoglu. Collaborative information retrieval: Toward a social informatics view of IR interaction. *Journal of the American Society for Information Science*, 49(12):1070–1080, 1998.

[7] M. Lesk. *Practical Digital Libraries: Books, Bytes, and Bucks.* Morgan Kaufmann, San Francisco, CA, 1997.

[8] S. Mizzaro. A concrete proposal for an automatically refereed scholarly electronic journal. Research Report UDMI/07/99, Department of Mathematics and Computer Science, University of Udine, Via delle Scienze, 206 — Loc. Rizzi — Udine, Italy, April 1999. Presented at a Workshop at ECDL'99 Conference, Paris.

[9] S. Mizzaro. An automatically refereed scholarly electronic journal. In C. Bavec, M. Gams, V. Rajkovič, I. Rozman, and J. Györkös, editors, *Information Society 2000 — Proceedings of the Third International Conference*, pages 121–126, Ljubljana, Slovenia, 17–19 October 2000. ISBN: 961 6303-26-0.

[10] Z. Nadasdy. A truly all-electronic journal: Let democracy replace peer review. *JEP*

the Journal of Electronic Publishing, 3(1), 1997. ISSN 1080-2711; http://www.press.umich.edu/jep/03-01/EJCBS.html.

[11] V. H. Resh. Science and communication: An author/editor/user's perspective on the transition from paper to electronic publishing. *Issues in Science and Technology Librarianship*, 19, 1998. ISSN 1092-1206; http://www.library.ucsb.edu/istl/98-summer/article3.html.

[12] D. S. Stodolsky. Consensus journals: Invitational journals based upon peer consensus. *Psycoloquy*, 1(15), 1990. Also appeared in: Datalogiske Skrifter (Writings on Computer Science), 29, 1990, Roskilde University Centre, Institute of Geography, Socioeconomic Analysis, and Computer Science (ISSN 0109-9779-29).

[13] Hal R. Varian. The future of electronic journals. *JEP the Journal of Electronic Publishing*, 4(1), 1998. ISSN 1080-2711; http://www.press.umich.edu/jep/04-01/varian.html.