ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-1614

Reputation in the Academic World

Nardine Osman and **Carles Sierra** [1]

Abstract. This paper proposes a computational model based on peer reviews for assessing the reputation of researchers and research work. We argue that by relying on peer opinions, we address some of the pitfalls of current approaches for calculating the reputation of authors and papers. We also introduce a much needed feature for review management: calculating the reputation of reviews and reviewers.

1 MOTIVATION

With open access gaining momentum, **open reviews** becomes a more persistent issue. Current institutional and multidisciplinary open access repositories lack the quantitative assessment of the hosted research items that will facilitate the process of selecting the most relevant and distinguished content. Common currently available metrics, such as number of visits and downloads, do not reflect the quality of a research product, which can only be assessed directly by peers offering their expert opinion together with quantitative ratings based on specific criteria. The articles published in the Frontiers book [1] highlight the need for open reviews.

To address this issue we develop an open peer review module, the Academic Reputation Model (ARM), as an overlay service to existing institutional or other repositories. The model calculates the reputation of authors, reviewers, papers, and reviews, by relying on peer opinions. We argue that this addresses some of the pitfalls of current approaches for calculating the reputation of authors and papers. It also introduces a much needed feature for review management, and that is calculating the reputation of reviews and reviewers.

2 ARM: ACADEMIC REPUTATION MODEL

2.1 Data and Notation

To compute its reputation measures, ARM requires a *reputation data set* (which should be extracted from existing repositories) specified as the tuple $\langle P, R, E, D, a, o, v\rangle$, where

- $P = \{p_i\}_{i\in\mathcal{P}}$ is a set of papers (e.g. DOIs).
- $R = \{r_j\}_{j\in\mathcal{R}}$ is a set of researcher names or identifiers (e.g. the ORCHID identifier).
- $E = \{e_i\}_{i\in\mathcal{E}} \cup \{\bot\}$ is a totally ordered evaluation space, where $e_i \in \mathbb{N} \setminus \{0\}$ and $e_i < e_j$ iff $i < j$ and $\bot$ stands for the absence of evaluation. We suggest the range [0,100], although any other range may be used, and the choice of range will not affect the performance.
- $D = \{d_k\}_{k\in\mathcal{K}}$ is a set of evaluation dimensions, such as *originality*, *technical soundness*, etc.
- $a : P \rightarrow 2^R$ is a function that gives the authors of a paper.
- $o : R \times P \times D \times Time \rightarrow E$, where $o(r, p, d, t) \in E$ is a function that gives the opinion of a reviewer, as a value in E, on a dimension d of a paper p at a given instant of time t.
- $v : R \times R \times P \times Time \rightarrow E$, where $v(r, r', p, t) = e$ is a function that gives the judgement of researcher r over the opinion of researcher r', on paper p as a value $e \in E$. Therefore, a judgement is a reviewer's opinion about another reviewer's opinion.

[1] Artificial Intelligence Research Institute (IIIA-CSIC), Barcelona, Spain, email: {nardine, sierra}@iiia.csic.es

2.2 Reputation of a Paper

The reputation of a paper is a weighted aggregation of its reviews, where the weight is the reputation of the reviewer (Section 2.4).

$$R_P(p) = \begin{cases} \dfrac{\sum_{\forall r\in rev(p)} R_R(r)\cdot o(r,p)}{\sum_{\forall r\in rev(p)} R_R(r)} & \text{if } |rev(p)| \geq k \\ \bot & \text{otherwise} \end{cases} \quad (1)$$

where $rev(p) = \{r \in R \mid o(r, p) \neq \bot\}$ denotes the reviewers of a given paper, and k is a parameter specifying the minimum number of reviews required for reputation to be calculates, and $\bot$ represents ignorance, that is, the reputation is not known.

2.3 Reputation of an Author

A researcher's author reputation is an aggregation of the reputation of her papers. The aggregation is based on the concept that *the impact of a paper's reputation on its authors' reputation is inversely proportional to the total number of its authors*. That is, if one researcher is the sole author of a paper, then this author is the only person responsible for this paper, and any (positive or negative) feedback about this paper is propagated as is to its sole author. However, if the researcher has co-authored the paper with several other researchers, then the impact (whether positive or negative) that this paper has on the researcher decreases with the increasing number of co-authors.

$$R_A(r) = \begin{cases} \dfrac{\sum_{\forall p\in pap(r)} \gamma(p)^\gamma \times R_P(p) + (1-\gamma(p)^\gamma)\times 50}{|pap(r)|} & \text{if } pap(r) \neq \emptyset \\ \bot & \text{otherwise} \end{cases} \quad (2)$$

where $pap(r) = \{p \in P \mid r \in a(p) \wedge R_P(p) \neq \bot\}$ denotes the papers authored by a given researcher r, $\bot$ describes ignorance, $\gamma(p) = \dfrac{1}{|a(p)|}$ is the coefficient that takes into consideration the number of authors of a paper (recall that $a(p)$ denotes the authors of a paper p), and γ is a tuning factor that controls the rate of decrease of the $\gamma(p)$ coefficient. Also note the multiplication by 50, which describes ignorance, as 50 is the median of the chosen range $[0, 100]$.

2.4 Reputation of a Reviewer

The reputation of a reviewer is an aggregation of the opinions over her reviews. We assume such opinions can be obtained, in a first instance, by other reviewers that *also reviewed the same paper*. However, as this is a new feature to be introduced in open access repositories and conference and journal paper management systems, we also provide an alternative: we take 'similarity' between reviews as a measure of the reviewers opinions about reviews. For instance, we assume that 'if my review is similar to yours then I may assume your judgement of my review would be good.' We note $v^*(r_i, r_j, p) \in E$ for the 'extended judgement' of r_i over r_j's opinion on paper p, and define it as an aggregation of opinions and similarities as follows:

$$v^*(r_i, r_j, p) = \begin{cases} v(r_i, r_j, p) & \text{if } v(r_i, r_j, p) \neq \perp \\ Sim(\bar{o}(r_i, p), \bar{o}(r_j, p)) & \text{If } \bar{o}(r_i, p) \neq \perp \text{ and } \bar{o}(r_j, p) \neq \perp \\ \perp & \text{Otherwise} \end{cases} \quad (3)$$

where Sim stands for an appropriate similarity measure. We say the similarity between two opinions is the difference between the two: $Sim(\bar{o}(r_i, p), \bar{o}(r_j, p)) = 100 - |\bar{o}(r_i, p) - \bar{o}(r_j, p)|$.

Now consider the set of judgements of r_i over reviews made by r_j as: $V^*(r_i, r_j) = \{v^*(r_i, r_j, p) \mid v(r_i, r_j, p) \neq \perp \text{ and } p \in P\}$. This set might be empty. We define the judgement of a reviewer over another one as a simple average:

$$R_R(r_i, r_j) = \begin{cases} \dfrac{\sum_{\forall v \in V^*(r_i, r_j)} v}{|V^*(r_i, r_j)|} & \text{if } V^*(r_i, r_j) \neq \emptyset \\ \perp & \text{otherwise} \end{cases} \quad (4)$$

Finally, the reputation of a reviewer r, $R_R(r)$, is an aggregation of judgements that her colleagues make about her reviews. We weight this with the reputation of the colleagues as a reviewer:

$$R_R(r) = \begin{cases} \dfrac{\sum_{\forall r_i \in R^*} R_R(r_i) \cdot R_R(r_i, r)}{\sum_{\forall r_i \in R^*} R_R(r_i)} & R^* \neq \emptyset \\ 50 & \text{otherwise} \end{cases} \quad (5)$$

where $R^* = \{r_i \in R \mid V^*(r_i, r) \neq \emptyset\}$. The default, representing ignorance, is 50 (as 50 is the median of the chosen range $[0, 100]$).

2.5 Reputation of a Review

The reputation of a review is an aggregation of its judgements, weighted by the reputation of their reviewers (Section 2.4).

$$R_O(r', p) = \begin{cases} \dfrac{\sum_{\forall r \in jud(r', p)} R_R(r) \cdot v^*(r, r', p)}{\sum_{\forall r \in jud(r', p)} R_R(r)} & \text{if } |jud(r', p)| \geq k \\ R_R(r') & \text{otherwise} \end{cases} \quad (6)$$

where $jud(r', p) = \{r \in R \mid v^*(r, r', p) \neq \perp\}$ denotes the set of judges of a review written by r' on paper p, and k is a parameter specifying the minimum number of judgements needed to calculate the reputation. The default is the reputation of the author of the review (her reputation as a reviewer).

3 EVALUATION

To evaluate the effectiveness of the proposed model, we have simulated a community of researchers, using NetLogo [3]. We simulated the creation of papers, reviews, and played with the parameters that tune the true quality of researchers (both as authors and reviewers), which impacts the true quality of their papers and reviews. The aim of the evaluation was to investigate how close are the calculated reputation values to the *true* values.

The results (Table 1) illustrate how the error of the reviewers' reputation increases as the number of bad reviewers increases in the community (where the reviewer's true quality is sampled from a beta distribution specified by the parameters α_R and β_R). The results also show how the error in the papers' reputation increases with the error in the reviewers' reputation, though at a smaller rate. One curious thing about these results is the constant error in the reputation of authors. To investigate this further, we played with the number of co-authors ($\#_{CA}$). Results (Table 2 show how an increasing number of co-authors increases the error in authors' reputation.

	Error in Reviewers' Reputation	**Error in Papers' Reputation**	**Error in Authors' Reputation**
α_R=5 & β_R=1	∼ 11 %	∼ 2 %	∼ 22 %
α_R=2 & β_R=1	∼ 23 %	∼ 5 %	∼ 23 %
α_R=1 & β_R=1	∼ 30 %	∼ 7 %	∼ 23 %
α_R=0.1 & β_R=0.1	∼ 34 %	∼ 5 %	∼ 22 %
α_R=1 & β_R=2	∼ 44 %	∼ 8 %	∼ 23 %
α_R=1 & β_R=2	∼ 60 %	∼ 9 %	∼ 20 %

Table 1: Impact of reviewers' quality on ARM's performance

	Error in Reviewers' Rep.		**Error in Papers' Rep.**		**Error in Authors' Rep.**	
	α_R=5; β_R=1	α_R=1; β_R=5	α_R=5; β_R=1	α_R=1; β_R=5	α_R=5; β_R=1	α_R=1; β_R=5
$\#_{CA}$=0	∼13%	∼54%	∼3%	∼9%	∼2%	∼7%
$\#_{CA}$=1	∼13%	∼57%	∼3%	∼9%	∼12%	∼15%
$\#_{CA}$=2	∼11%	∼60%	∼2%	∼9%	∼22%	∼20%

Table 2: Impact of number of co-authors on authors' reputation

4 CONCLUSION

We have presented the ARM reputation model for the academic world that calculates the reputation of researchers, both as authors and reviewers, and their research work, as well as the reputation of reviews. The model is based on peer reviews. For further details we refer the interested reader to [2].

ACKNOWLEDGEMENTS

This work has been supported by the projects CollectiveMind (funded by MINECO, grant # TEC2013-49430-EXP), and Open Peer Review Module for Repositories (funded by OpenAIRE).

REFERENCES

[1] Nikolaus Kriegeskorte and Diana Deca, eds. *Beyond open access: visions for open evaluation of scientific papers by post-publication peer review*, Front. Comput. Neurosci., 2012.

[2] Nardine Osman and Carles Sierra, 'Reputation in the academic world', in *Proc. of Trust in Agent Societies*, eds., J. Zhang, R. Cohen, and M. Sensoy, volume 1578, pp. 1–17. CEUR-WS.org, (2016).

[3] Seth Tisue and Uri Wilensky, 'Netlogo: Design and implementation of a multi-agent modeling environment', in *Proc. of Agent2004*, pp. 161–184, (2004).