Characterizing Contribution Strategies and Quality with Bi-Partite Networks in Wikipedia

Maximilian Klein
OCLC Research
777 Mariners Island Blvd
San Mateo, CA, 94404
kleinm@oclc.org

Thomas Maillart
School of Information
University of California,
Berkeley, 102 South Hall
Berkeley, CA 94720
thomas.maillart
@ischool.berkeley.edu

John Chuang
School of Information
University of California,
Berkeley, 102 South Hall
Berkeley, CA 94720
chuang
@ischool.berkeley.edu

ABSTRACT
In a world of collaboration, understanding how individuals achiev We introduce and test a recursive algorithm toalgorithm to bi-partite networks of relations between two kinds of nodes (here, editors making changes to articles): the expertise of editors is assessed from the quantity and quality of articles they have edited. Conversely, the quality of an article depends on the number and the expertise of editors who have modified the article. And so on recursively. While wikiRanks incorporates only bi-partite links input information, we find high rank correlations ($\rho = 0.7 \pm 0.1$) with usual Wikipedia editors' expertise and articles' quality metrics. The wikiRanks algorithm also provides deep insights on the structure of online collaboration. We find that editors in some categories of Wikipedia articles achieve more quality with a large number of editors per article, while for other categories, quality is more achieved as a result of expertise of editors.The algorithm we have developed to assess and understand individual contributions to Wikipedia, can be applied to any collaborative environment.

Author Keywords
open collaboration, bi-partite networks, contribution performance

INTRODUCTION
In open collaboration, knowledge is produced by a multitude of contributors, according to the rules of peer-production as a form of labor organization emerging in computer and Internet supported environments [1]. The most basic component of peer-production is *task self-selection*: participants in open collaboration freely choose how and when to contribute, with limited or no vertical organization. Starting with open source software development, open collaboration has permeated in a broad variety of industries [2]. Wikipedia is one of the most successful examples of open collaboration, going head-to-head with major online content providers, and competing on accuracy with traditionally edited encyclopedia [6].

Nonetheless, it remains hard to understand how knowledge is pulled from the self-organization of many contributing individuals with their heterogenous backgrounds and motivations. The production process of many open collaboration communities exhibits complex critical cascades of input, which in turn lead to super-linear productive bursts of contributions [17]. In other words, the dynamics of contributions remain deeply entangled, and therefore, tell little on how the heterogenous inputs by individuals in open collaboration are organized to achieve best quality. Moreover, the multiplicity of contribution organization in open collaboration suggests that community organization is "tuned" to a specific kind of knowledge.

Here, we propose that the contribution structure is deeply tied to type of knowledge produced. We develop an iterative approach to account for the complex relationships (i.e. the bi-partite network) between contributors and units of knowledge, such as articles in Wikipedia. For 12 Wikipedia categories, we demonstrate how the structure of knowledge contribution can be disentangled, and how it influences the quality of knowledge produced.

The paper is organized as follows. The reader is first introduced to relevant literature . The method, data employed and results are then presented and discussed. We finally present future research directions and conclude.

RELATED WORK
To the best of our knowledge, our paper is the first attempt to quantify the structure and the impact of collaboration on value creation in open collaboration environments. Our model of bi-partite random walkers follows in the lineage of bi-partite networks – with two node types – in the global economy of countries competing for exporting products [10, 9]. The proposed *method of reflections* helps understand the competitive advantage (i.e., *fitness*) of countries from the types of products they sell, and moreover, whether other countries export similar products *ubiquity*. Most competitive countries are found to export non-ubiquitous products, for which they can charge higher price. The reflexive model has been iteratively improved and complemented in more recent work, mainly to fix the robustness of the method

CHI'12, May 5–10, 2012, Austin, Texas, USA.

of reflections.[18, 5, 19, 4]. Caldarelli et al. [3] have proposed an alternative method, based on biased Markov chains, which helps further understand the relations between two types of nodes in bi-partite networks.

At the opposite of networks of economic competition, collaboration networks have been studied early on in network sciences, in particular networks of co-authorship in scientific publications [13], as well as patterns of self-organization in bi-partite networks of actors-movies[15]. Similarly, the analysis of patterns in the Wikipedia bi-partite networks confirmed the existence overlapping cliques of densely connected articles and editors [11]. In the same study, clustering of densely connected cliques into larger modules [7] showed that topics aggregate editors by interest with highly coordinated efforts in densely populated clusters [11].

In a series of Matlab contests aimed at collectively solving NP-hard problems, it was found that work shared as a public good helps individuals quickly reuse existing results, and thus, find better algorithms [8]. Efforts have also been undertaken to assess the *expertise* of contributors [**?**], as well as the quality of articles [20] in Wikipedia.

METHOD

Measuring the structure of value creation by individuals is nearly impossible in most in open collaboration projects, in particular when they do not involve writing software code, which can be compiled, executed and tested on computers. However, like in Wikipedia, the most common way to code knowledge is natural language, which cannot be executed and tested for performance. Bi-partite networks models of editors and their contributions to article provide a very general framework to understand the structure of value creation in open collaboration.

We consider a binary matrix $M_{ea} \equiv M$ of editors having contributed to a category of articles on Wikipedia: M_{ea} takes value 1 if editor e has modified article a, 0 otherwise. For the category *Feminist Writers*, as presented on Figure 1, M exhibits a triangular structure with a few editors having modified a lot of articles while many editors have contributed to at best a few articles. Similarly, a few articles were modified by a lot of editors, while a vast majority of articles were touched only by one or very few authors. M is the only input of the *bi-partite random walker* model described thereafter, for characterizing the structure and the value of contributions in open collaboration.

Given M, a simple way to assess the contribution value of an editor is obtained by summing the number of articles ever edited over a set of articles. Similarly, a simple *quality* measure for an article is the sum of editors who have ever modified it, following the famous adage on open source development: “Given eyeballs, all bugs shallow” [16]. These crude expertise and quality metrics for respectively editors and articles given by,

$$\begin{cases} u_e^{(0)} = \sum_{a=1}^{N_a} M \equiv k_e \\ u_a^{(0)} = \sum_{e=1}^{N_e} M \equiv k_a \end{cases} \tag{1}$$

are the zeroth order of our algorithm. They also constitute the initial step of the *method of reflections* proposed by Hidalgo et al. [10, 9], which derives the value of producing entities (i.e. editors) from products (i.e. articles), and *vice versa*, iteratively. Adapted to Wikipedia, the first steps help capture the intuition behind the method of reflections for open collaboration:

- **1$^{\text{st}}$ order iteration** : if an article has been edited by higher expertise editors (i.e. who have contributed to higher value articles), can we infer its quality? Conversely, if an editor has contributed to higher quality articles (i.e. who have been changed by more expert editors), can we evaluate the editor's expertise ?
- **2$^{\text{nd}}$ order iteration** : if an article has been changed by higher expertise editors who have edited higher value articles, which in turn have been edited by higher expertise contributors, can we infer its quality? Conversely, if an editor has edited higher quality articles, which have been edited by better editors who have edited higher quality articles, can we evaluate the expertise of the editor?
- **and so on, recursively**.

Although the interpretation is difficult to imagine passed the very first iteration steps, the algorithm incorporates additional information on the quality (resp. expertise) of the article (resp. editor) from the previous iteration from expertise (resp. value) of editors (resp. of articles). The higher order steps of the method of reflection writes,

$$\begin{cases} u_e^{(n+1)} = \frac{1}{k_e} \sum_{a=1}^{N_a} M\, u_a^n \\ u_a^{(n+1)} = \frac{1}{k_a} \sum_{e=1}^{N_e} M\, u_e^n. \end{cases} \tag{2}$$

The method of reflection however suffers a number of problems, which are mostly rooted in the equal weights given to mutual influences of editors expertise over articles quality, and of articles quality on editors expertise. As a result, the method converges to a fixed point with all editors (resp. all articles) having the same value, as a case of consensus dynamics [3]. It has therefore been proposed to consider the *reflective method* as a Markov process of random walkers on a bi-partite network, jumping with some probability from one node type to another node type [3]. A schematic representation of the random walk process on a bi-partite network is depicted on Figure 2. The intuition is the following: a random walker jumps with some probability from an editor to a given article (i.e. the editor's expertise is positively influenced by the article's quality), and with another probability from an article to a given editor (i.e. the value of the article is positively by the editor's expertise). The matrix M determines the probability functions for each node. If two nodes e and a are not

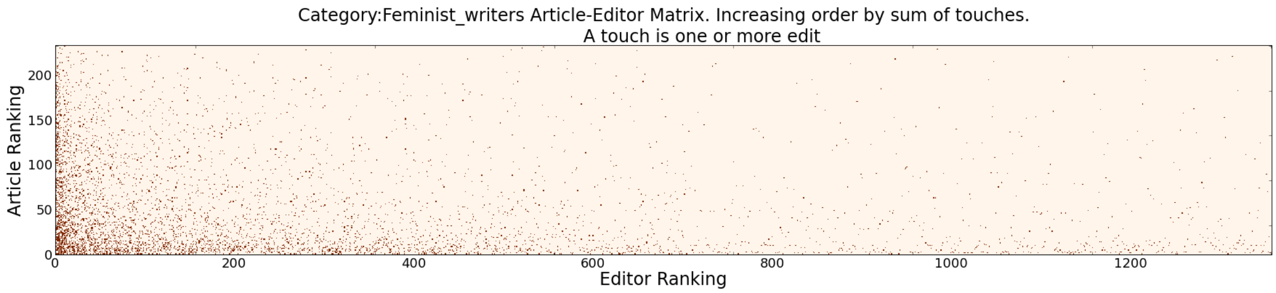

Figure 1. Typical M matrix for a Wikipedia category (here, *Feminist Writers*) ordered on both dimensions by descending order of number of articles modified by an editor (horizontal axis) and of number editors who have modified an article (vertical axis). The structure of M is triangular and shows that some editors have a pervasive activity over articles, while most editors edit only a few. Similarly, some articles receive widespread attention by editors, while most articles are modified only by a few editors.

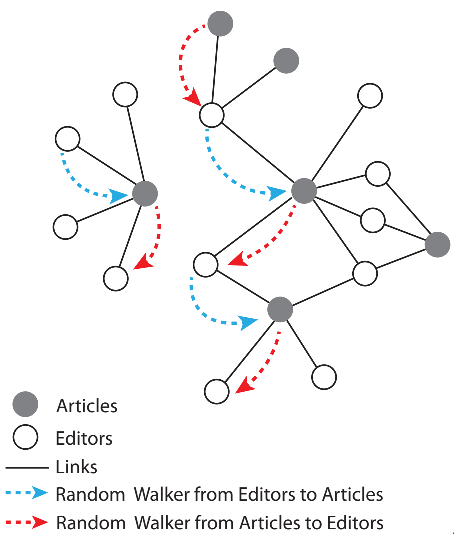

Figure 2. Representation of random walkers jumping from editors to articles (red dotted arrows) and from articles to editors (blue dotted arrows). The intuition is the following: a random walker jumps with some probability from an editor to a given article (i.e. the editor's expertise is positively influenced by the article's quality), and with another probability from an article to a given editor (i.e. the value of the article positively influences the editor's expertise).

directly connected $M_{ea} = 0$, and the transition probability is 0. The *bi-partite network random walker* model is an extension of the *Page Rank* Google search algorithm [14] to two kind of nodes

According to [3], we reformulate the method of reflection to account for jumps of the random walker on the bi-partite network of editors and articles. Let's call $w_e^{(n)}$ the expertise of an editor and $w_a^{(n)}$ the quality of an article at the n^{th} iteration. We define the following Markov process on the bi-partite network of collaboration,

$$\begin{cases} w_e^{(n+1)}(\alpha,\beta) = \sum_{a=1}^{N_a} G_{ea}(\beta)\, w_a^{(n)}(\alpha,\beta) \\ w_a^{(n+1)}(\alpha,\beta) = \sum_{e=1}^{N_e} G_{ae}(\alpha)\, w_e^{(n)}(\alpha,\beta) \end{cases} \tag{3}$$

with G_{ea} the probability to jump from article a to editor e in a single step, and the probability G_{ae} to jump from editor e to article a also in a single step. These transition probabilities are given by,

$$\begin{cases} G_{ea}(\beta) = \frac{M_{ea} k_e^{-\beta}}{\sum_{e'=1}^{N_e} M_{e'a} k_{e'}^{-\beta}} \\ G_{ae}(\alpha) = \frac{M_{ea} k_a^{-\alpha}}{\sum_{a'=1}^{N_a} M_{ea'} k_{a'}^{-\alpha}} \end{cases} \tag{4}$$

The transition matrices $G_{ea}(\beta)$ and $G_{ae}(\alpha)$ depend only on the initial conditions: the binary matrix M, as well as k_e and k_a given by (1). The variables α and β control the transition probabilities in a similar way. We shall therefore explain only how β influences the probability to jump from an article to an editor (i.e. the value of the article positively influences the editor's expertise). For $\beta = 0$, we recover the zero$^{\text{th}}$ order iteration (1). For $\beta > 0$, the probability to jump from article a to editor e is a power law function $\sim 1/k_e^{\beta}$ of the sum of articles k_e modified by editor e. Hence, the larger k_e, the lower the probability to jump from a to e relative to other editors. On the contrary, if $\beta < 0$ the probability to jump from an article to an editor is a positive function of the sum of articles modified by the editor. For $-1 < \beta < 0$, the function is concave, while for $\beta < -1$, the function is convex, which means that the more articles have been edited by the editor, the even more the positive influence on articles. In a nutshell, β relates the amount of articles edited on the overall editor's expertise, which in turn has an influence on each edited article (along with the influence of other editors). If $\beta >> 0$ is positive, the positive influence of the number of contributed articles on the editor's expertise decreases. If β close to 0, the number of contributed articles increase linearly the editor's expertise. The same considerations hold for α and the probability $G_{ae}(\alpha)$ to jump from an editor to an article (i.e. the expertise of the editor positively influences the quality of

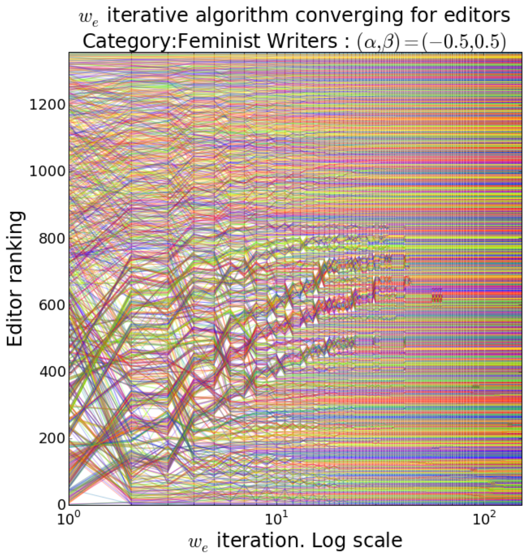

Figure 3. Convergence of the ranked expertise w_e of editors having contributed to articles in the Feminist Writers category on Wikipedia for arbitrary control parameters: $(\alpha, \beta) = (-0.5, 0.5)$. Starting from the sum of contributed articles as the initial step, we can see how the algorithm progressively ranks editors: some editors with initial lowest rank, i.e. with few articles edited, get a higher rank as the number of iterations increases. Similarly, some initially high ranked editors, gradually drop in the ranking. In the case Wikipedia categories, the algorithm converges with clearly stable ranks, after nearly 50 iterations.

an article).

We have verified that the *bi-partite network random walker* method converges. Figure 3 shows the evolution expertise w_e ranked among editors having contributed to articles in the *Feminist Writers* category on Wikipedia for arbitrary control parameters: $(\alpha, \beta) = (-0.5, 0.5)$. Starting from the sum of contributed articles as the initial step, we can see how the algorithm progressively ranks editors: some editors with initial lowest rank, i.e. with few articles edited, get a higher rank as the number of iterations increases. Similarly, some initially high ranked editors, gradually drop in the ranking. In the case Wikipedia categories, the algorithm converges with clearly stable ranks, after nearly 50 iterations.

The parameters α and β control how editors and articles influence each others at the level of the whole bi-partite collaboration network. Upon calibration they also inform on the structure of peer-production, and how contributions benefit to the whole open collaboration project. Whenever calibration shows that the evolution of α and β can be predicted, then editors' expertise and articles' quality can also be predicted up to statistical errors. Here, we demonstrate the calibration steps forα and β and how the control parameters inform on the structure of value creation in open collaboration. We also document the evolution of these parameters as Wikipedia categories get increasingly enriched with new contributions.

DATA

We aim to calibrate α and β for 12 categories of articles on Wikipedia (see Table 1), as well as their evolution over 13

Category	Articles	Editors	Edits
American male novelists	2,460	9,946	224,783
2013 films	1,896	5,215	150,956
American women novelists	1,936	5,968	138,716
Nobel Peace Prize laureates	104	4,165	91,522
Sexual acts	93	2,190	45,901
Economic theories	212	1,145	28,658
Feminist writers	233	1,357	25,738
Yoga	123	730	25,315
Military history of the US	180	854	20,172
Counterculture festivals	66	578	10,515
Computability theory	92	272	7,117
Bicycle parts	70	210	4,981

Table 1. Size statistics of investigated Wikipedia categories sorted by total edits.

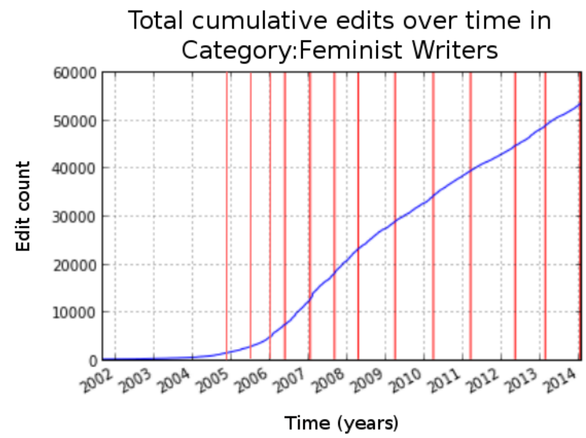

Figure 4. Cumulative edits made in Category *Feminist writers* (blue line). Vertical red lines represent the 13 snapshots taken at 2.5%, 5%, 7.5% and then, 10%, 20%, 30%, ... , 100% of edits.

snapshots as presented on Figure 4. For each category and snapshot, we built the binary matrix M by parsing all edit histories of all articles up to the snapshot time. In order to eliminate page vandals, we considered that an editor made a contribution to an article only if she has made 5 or more edits: we set $M_{ae} = 1$ for editor e having modified article a, and $M_{ae} = 0$ otherwise. We also discarded all software robots *Bots* that programmatically edit Wikipedia.

To calibrate α and β, we resorted to state-of-the-art exogenous evaluations for editor expertise $\bar{w}_e$ and article quality $\bar{w}_a$. From the exogenous evaluations, we ranked editors and articles according to their expertise and quality respectively. We then performed a grid search for values of α^* and β^*, which maximize the Spearman rank-correlation ρ_e and ρ_a between rankings obtained from the *bi-partitite random walker* model (w_e, w_a) and from exogenous metrics $(w_a, \bar{w}_a)$. Actually, (α^*, β^*) must maximize both ρ_e and ρ_a, even though rho_e and ρ_a might actually be different. The optimization function of (α^*, β^*) if given by,

$$\begin{cases}(\alpha^*, \beta^*) = argmax(\rho_e) \\ (\alpha^*, \beta^*) = argmax(\rho_a).\end{cases} \tag{5}$$

(α^*, β^*) characterize how value flow from editors to articles, and conversely, in the bi-partite network of collaboration in Wikipedia.

Editors expertise is best represented by labor hours [**?**], and was calculated for each editor by taking contribution history up to the snapshot point. All edits made within 1 hour of a previous edit are counted in an *edit session*. If more than one hour separates two edits, a new edit session starts. The expertise expressed in labor hours is the sum of edit sessions. In order to capture expertise from a specific field (instead of Wikipedia editing expertise), we only consider edits for a given category. The measure of article quality is performed through a combinations of 5 text analysis metrics: (i) ratio of mark-up to readable text, (ii) number of headings, (iii) article length, (iv) citations per article length, (v) outgoing intra-Wiki links. **[explain the remix version reused from [12]]**.vWe performed a Principal Component Analysis (PCA) for each category and snapshot, in order o reduce dimensionality from 5 metrics to a single one (i.e. the principal component). The variance explained by the principal component varied between 0.5 and 0.72.

RESULTS

To understand how value of contributions *flows* from editors to articles, and from articles to editors, we have calibrated the control parameters (α^*, β^*) of the *bi-partite network random walker* on 12 Wikipedia categories (c.f. Table 1) with 13 snapshots each (Figure 4). Figure 5 shows typical optimization landscapes, which maximize the rank correlation ρ_e (resp. ρ_e) between editors expertise w_e (resp. articles value) obtained from the model and expertise obtained from state-of-the-art measures $\bar{w}_e$ (resp. ρ_a).

The maximum achievable rank-correlation with ground-truth expertise metrics for editors [**?**] and quality for articles [**?**] shows that the bi-partite model of influence accounts particularly well for both quality of articles ($0.58 < \rho_a < 0.91$) and expertise of editors ($0.46 < \rho_e < 0.75$) at the last snapshot. Actually, the *bi-partite network random walker* model reproduces very well early the ranking of editors and articles according to the ground-truth metrics (see Figure 6). In particular, articles quality is very well accounted for, while the level of correlation with the ground-truth of editors expertise exhibits a rather linear, or at least slightly concave increase.

For the latest snapshot (i.e., the state contributions in February 2014), we find that the best possible α^* is 0 in all circumstances, and β^* varies across categories. Table 2 shows the categories ordered by the optimal β^* (and $\alpha^* = 0$ for the sake of clarity), as well as the corresponding maximum rank correlations ρ_e and ρ_a. Since there is no single optimal value for (α^*, β^*) , but rather a space of optimal values, we have searched for a value which maximizes both ρ_e and ρ_a. $\alpha^* = 0$ means that editors expertise always benefits from contributions as a linear function of the number of articles edited (and so on at higher orders). However, β^* exhibits a continuum of values between 0 (*Bicycle parts* and *US Military History*) and 1.52 (*Sexual Acts"*). β controls the influence of the number of editors on the quality of a given article (and so on at higher orders). When $\beta \approx 0$, the quality of articles is increased as

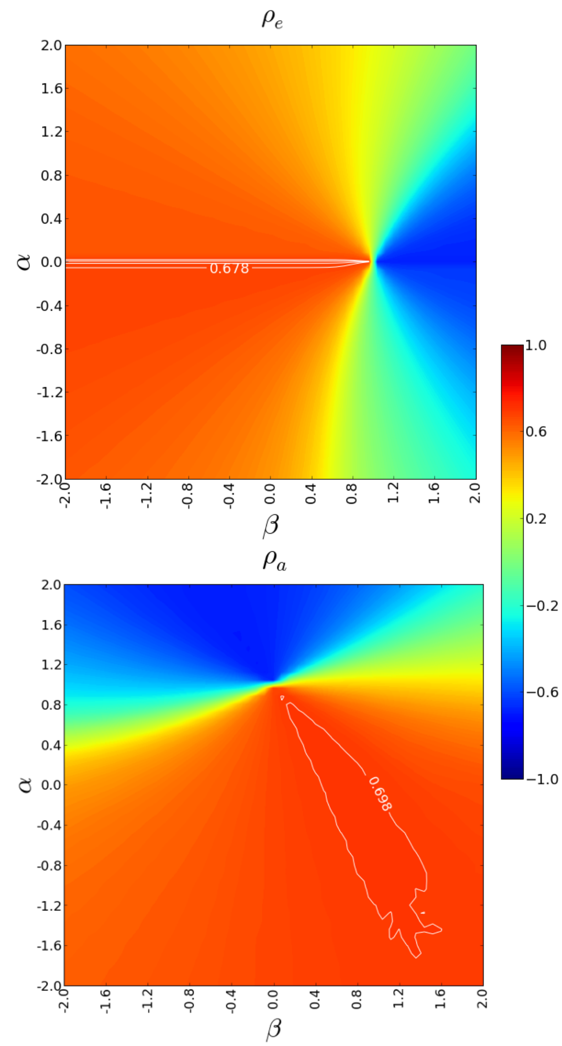

Figure 5. Typical landscape of maximum correlation as a function of α and β for articles (upper panel) and editors (lower panel). The contour line shows the 95th percentile of the rank correlation over the landscape. which category is represented here? Even though we don't see it here, there is a sharp transition at $\alpha = 0$ with high correlation for $\alpha < 0$ and low correlation for $\alpha > 0$. Whether α defined is unclear, but doubt it

a linear function of the number of editors who have modified them. For $\beta \gg 0$, the marginal gain of having more editors for a given article decreases.

β is found generally stable over snapshots for categories, past the first 10% of overall contributions. The stability of (α^*, β^*) confirm that the control parameters of the *bi-partite network random walker* describe a robust feature of the structure of value creation, at least in Wikipedia. This additional result is also a first step towards predictions of editors and articles, given the input made by editors.

DISCUSSION

Building the method of reflection previously used for global economic networks of production, we have applied and tested the *bi-partite network random walker* model in the context

	Category	ρ_a	ρ_e	α	β
1	Bicycle parts	0.90	0.46	0.00	0.00
2	Military history of the US	0.58	0.70	0.00	0.00
3	Computability theory	0.77	0.56	0.00	0.32
4	American male novelists	0.67	0.75	0.00	0.40
5	2013 films	0.72	0.55	0.00	0.48
6	Economic theories	0.74	0.70	0.00	0.48
7	American women novelists	0.63	0.75	0.00	0.64
8	Feminist writers	0.70	0.69	0.00	0.72
9	Yoga	0.64	0.57	0.00	1.12
10	Nobel Peace Prize laureates	0.91	0.66	0.00	1.20
11	Counterculture festivals	0.80	0.61	0.00	1.36
12	Sexual acts	0.63	0.66	0.00	1.52

Table 2. Categories ordered by increasing β obtained from best rank-correlation ρ_a and ρ_e of the *bi-partite network random walker* with the ground truth. As shown on the upper panel of Figure 5, highest rank-correlation is always obtained for $\alpha = 0$ suggesting that articles get systematically enriched from more editors. Different values of β, however, show that some categories require more specialized authors (β large), while others

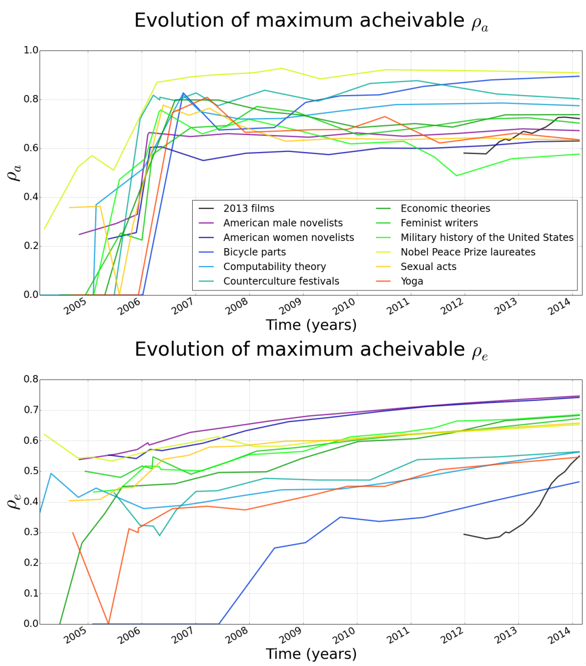

Figure 6. Evolution of Spearman ρ rank correlations between the ranking obtained from the calibrated model and the grand-truth values for each category and for editors (upper panel) and articles (lower panel). The correlations are generally quite high : $0.46 < \rho_e < 0.75$ with $\langle \rho_e \rangle = 0.64$ for editors and $0.57 < \rho_a < 0.91$ with $\langle \rho_a \rangle = 0.72$. ρ_a is stable over time, which means that the quality of articles can be well captured early on by the model. However, ρ_e exhibits a convex increase over time, suggesting that it takes time (i.e., lots of edits) to capture well the expertise of editors. the panel is weirdly cut on the left. It would be great to see when is the first article (signaled e.g. by a star) and the first snapshot signal by a vertical cross ("+"). All snapshot points should be mentioned by a cross

of Wikipedia open collaboration. Our results show that the model accounts well for the quality of articles ρ_a around 0.64 and for the expertise of contributors ρ_e around 0.72, even though the ground truth quality and expertise metrics are very rough. Moreover, the evolution of ρ_e and ρ_a as categories of articles get further contributed, show astonishing stability. In particular, the adequacy of the model to article quality is very high early on, followed by stationarity, hence suggesting that the model can quickly catch the quality of articles. For editors expertise, the adequacy increases steadily as categories get further developed.

This difference might be due to the roughness of the grand truth for editors v_e, expressed in labor-hours, compared to the quality of v_a, which is an aggregate measure of five precise quality metrics. Nevertheless, the correlation of editors' ranking with v_e increases: ρ_e exhibits a convex increase over time, suggesting that it takes time (i.e., lots of edits) to capture well the expertise of editors as well as for editors. This is striking because the method is reflexive and the same information is incorporated on both dimensions from the input matrix $\mathbf{M}$. While we have no definitive answer, we interpret this result in the following way: from Figure **??**, we see that there are usually more editors than articles for a category. This means that the probability for an article to get contributions early on is higher than the probability to find editors who have contributed to a lot of articles early. In other words, it takes more time to rightly rank editors because there are more compared to the number of articles in a given category.

We have also found $\alpha \approx 0$ for all categories, reflecting the positive influence of the number of articles edited on editors expertise. This stylized fact tells actually a lot about the model. The ground-truth metric for editors expertise takes only into consideration the labor hours [**?**], and the more time is spent editing a category, the more likely the editor will have modified a large quantity of articles. The value of α can therefore be explained entirely by the nature of the ground-truth metric. Although it would require further testing with a broad variety of metrics, it seems that the bi-partite network random walker model can *tune* to whatever ground-truth metric used for calibration. In other words, the values of α and β only reflect the structure of value creation in open collaboration, *given* the chosen ground-truth metrics.

Finally we find that whether article ubiquity is a measure of quality or not is category dependent. With these two examples we find footing to propose to take our β coefficient to be a proxy for the "collaborativeness," of a category. Max : please explain more what you had in mind here

Overall, it is amazing amazing that so little information (binary matrix M_{ae} with information only on whether an editor as modified a given article) provided as an input to the algorithm is sufficient to achieve a high level of fidelity with ground-truth metrics. It suggests some kind of *less is more* mechanism, which has implications on the computational aspects of the algorithm.

LIMITATIONS AND FUTURE WORK

The results we have presented show that contributions – and the value they provide to the collective good – can be disentangled, with the bi-partite network random walker model. However, these results suffer from a number of limitations, which call for further work and testing.

In our model, the value of contributions is only inferred from the binary matrix of contributions.

less is more

According to the original philosophy of the reflection method [10], the capabilities of producing entities (e.g., countries) that constrain the production are not observable. We believe this is feasible in the case open collaboration. We can for instance identify what an editor do best to improve an article, among the five metrics (ratio of mark-up to readable text, number of headings, article length, citations per article length, and outgoing intrawiki links) we have used to assess the quality of an article.

We believe there are two further directions to improve our results. First we have followed the philosophy of the reflection method and we have used rankings by defaults. However, the bi-partite network random walker method provides absolute values, which might have a meaning in the context of open collaboration. Second, we took the very simplest information for the input matrix $\mathbf{M}$ (i.e. whether an editor has modified an article). We wonder how the performance of the method might change if richer information is incorporated in the matrix (e.g., number of edits). Would it improve performance or on the contrary reduce it ?

- Improve search algorithm for calibration of α and β. For the time being we are performing grid search, which is less than optimal.

- Although the results are robust given the poor input (whether an editor has contributed to an article or not) Change the input : what happens if a richer input

- we shall explore ways to improve the correlation and the predictive power of the algorithm. Let's consider that we are able to successfully calibrate α and β and we know how they are likely to evolve as a category of articles gets further enriched, then as editors contribute, it will be possible to know how they have created added value to the collective project.

CONCLUSION

Endogenous expertise acquisition by editors and quality of Wikipedia articles is a critical problem to overcome in order to understand how people, as individuals, contribute to the achievement of high quality knowledge as public good and how articles greatly benefit from their iterative inputs. For 12 categories of Wikipedia articles, we have identified and calibrated metrics from a bi-partite network random walker model. These metrics quantify how editors expertise and articles quality is derived from the variety of inputs made by individuals over time. These metrics give direct insights on how some categories of Wikipedia articles mainly benefit from either a broad set of small inputs by lay-persons, or on the contrary, from a few "experts" in the field. They also tell whether editors expertise is rather achieved through contributions to a broad set of articles, or by concentrating on a few high level articles. While this method suffers limitations, a similar approach can be extended for better understanding how contributions are valued in open collaboration. Similarly, it could applied to a broad variety of situations, such as social coding (e.g., Github) , or collaborative rating (e.g., Amazon or Yelp reviews), to help understand to origins of quality and reliability, which is often highly recognized and valued in open collaboration.

ACKNOWLEDGMENTS

One of the authors (T.M.) acknowledges support from the Swiss National Science Foundation (Grant Nr. PA00P2-145368).

REFERENCES

1. Benkler, Y. Coase's Penguin, or, Linux and "The Nature of the Firm". *The Yale Law Journal 112*, 3 (Dec. 2002), 369+.
2. Benkler, Y. *The Penguin and the Leviathan: How Cooperation Triumphs over Self-Interest*, 1 ed. Crown Business, Aug. 2011.
3. Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., and Tacchella, A. A network analysis of countries' export flows: firm grounds for the building blocks of the economy. *PloS one 7*, 10 (2012).
4. Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., and Pietronero, L. Measuring the Intangibles: A Metrics for the Economic Complexity of Countries and Products. *PloS one 8*, 8 (2013).
5. Cristelli, M., Tacchella, A., Gabrielli, A., Pietronero, L., Scala, A., and Caldarelli, G. Competitors' communities and taxonomy of products according to export fluxes. *The European Physical Journal-Special Topics 212*, 1 (2012).
6. Giles, J. Internet encyclopaedias go head to head. *Nature 438*, 7070 (Dec. 2005), 900–901.
7. Guimerà, R., Sales-Pardo, M., and Amaral, L. A. Module identification in bipartite and directed networks. *Physical Review E 76*, 3 (2007), 036102.
8. Gulley, N., and Lakhani, K. R. The Determinants of Individual Performance and Collective Value in Private-Collective Software Innovation. *Social Science Research Network Working Paper Series* (Feb. 2010).
9. Hidalgo, C. A., and Hausmann, R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences 106*, 26 (June 2009), 10570–10575.
10. Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. The Product Space Conditions the Development of Nations. *Science 317*, 5837 (July 2007), 482–487.
11. Jesus, R., Schwartz, M., and Lehmann, S. Bipartite Networks of Wikipedia's Articles and Authors: A Meso-level Approach. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*, WikiSym '09, ACM (New York, NY, USA, 2009), 1–10.
12. Klein, M. Kumusha Takes Wiki: Actionable Metrics for Uganda and Côte d'Ivoire, 2014.

13. Newman, M. E. J. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences 98*, 2 (Jan. 2001), 404–409.
14. Page, L., Brin, S., Motwani, R., and Winograd, T. The PageRank citation ranking: Bringing order to the web.
15. Ramasco, J. J., Dorogovtsev, S. N., and Pastor-Satorras, R. Self-organization of collaboration networks. *Physical review E 70*, 3 (2004), 036106.
16. Raymond, E. The cathedral and the bazaar. *Knowledge, Technology & Policy 12*, 3 (Sept. 1999), 23–49.
17. Sornette, D., Maillart, T., and Ghezzi, G. How Much is the Whole Really More than the Sum of its Parts? 1 + 1 = 2.5: Superlinear Productivity in Collective Group Actions, May 2014.
18. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. A new metrics for countries' fitness and products' complexity. *Scientific reports 2* (2012).
19. Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. Economic complexity: Conceptual grounding of a new metrics for global competitiveness. *Journal of Economic Dynamics and Control* (2013).
20. Wang, M. W., Cosley, D., and Riedl, J. Tell Me More: An Actionable Quality Model for Wikipedia. In *Proceedings of the 9th International Symposium on Open Collaboration*, WikiSym '13, ACM (New York, NY, USA, 2013).