A Game Theoretic Analysis of Collaboration in Wikipedia

S. Anand[*], O. Arazy[†], N. B. Mandayam[*] and O. Nov[‡]

[*] Wireless Information Networks Laboratory (WINLAB), Rutgers University
E-mail: {anands72,narayan}@winlab.rutgers.edu
[†] Alberta School of Business, Edmonton
E-mail: ofer.arazy@ualberta.ca
[‡] Department of Technology Management and Innovation,
Polytechnic Institute of New York University
E-mail: on272@nyu.edu

Abstract. Peer production projects such as Wikipedia or open-source software development allow volunteers to collectively create knowledge-based products. The inclusive nature of such projects poses difficult challenges for ensuring trustworthiness and combating vandalism. Prior studies in the area deal with descriptive aspects of peer production, failing to capture the idea that while contributors collaborate, they also compete for status in the community and for imposing their views on the product. In this paper, we investigate collaborative authoring in Wikipedia, where contributors append and overwrite previous contributions to a page. We assume that a contributor's goal is to maximize ownership of content sections, such that content owned (i.e. originated) by her survived the most recent revision of the page. We model contributors' interactions to increase their content ownership as a non-cooperative game, where a player's utility is associated with content owned and cost is a function of effort expended. Our results capture several real-life aspects of contributors interactions within peer-production projects. Namely, we show that at the Nash equilibrium there is an inverse relationship between the effort required to make a contribution and the survival of a contributor's content. In other words, majority of the content that survives is necessarily contributed by experts who expend relatively less effort than non-experts. An empirical analysis of Wikipedia articles provides support for our model's predictions. Implications for research and practice are discussed in the context of trustworthy collaboration as well as vandalism.

Index Terms-Peer production, Wikipedia, collaboration, non-cooperative game, trustworthy collaboration, vandalism

1 Introduction

Recent years have seen the emergence of a web-based peer-production model for collaborative work, whereby large numbers of individuals co-create knowledge-based goods, such as Wikipedia, and open source software [30], [12], [23], [14], [18], [24], [19]. Increasingly, individuals, companies, government agencies and other organizations rely on peer-produced products, stressing on the need to ensure trust worthiness of collaboration (e.g., deter vandalism) as well as the quality of end products.

Our focus in this study is Wikipedia, probably the most prominent example of peer-production. Wikipedia has become one of the most popular information sources on the web, and the quality of Wikipedia articles has been the topic of recent public debates. Wikipedia is based on wiki technology. Wiki is a web-based collaborative authoring tool that allows contributors to add new content, append existing content, or delete and overwrite prior contributions. Wikis track the history of revisions similarly to version control systems used in software development allowing users to revert a wiki page to an earlier revision [25], [36], [15].

Peer production projects face a key tension between inclusiveness and quality assurance. While such projects need to draw in a large group of contributors in order to leverage "the wisdom of the crowd," there is also requirement for accountability, security, and quality control [17], [20], [35]. Quality assurance measures are necessary not only in cases of vandalism; conflicts between contributors could also result from competition. For example, contributors to Wikipedia may wrestle to impose their own viewpoints on an article especially for controversial topics or attempt to dominate subsets of the peer-produced product. Another example is when contributors seeking status within the community compete to make the largest contribution, and in the process overwrite others' previous contributions. The result of such competitions is often "edit wars" where articles are changed back-and-forth between the same contributors.

Prior studies investigating an individual's motivation for contributing content to Wikipedia have identified a large number of motives driving participation[12],[22], including motives that are competitive in nature, such as ego, reputation enhancement, and the expression of one's opinions [22], [26]. However, studies investigating individuals did not consider the competitive dynamics emerging from motives such as reputation. Research into group interactions at Wikipedia, have tended to emphasize the collaborative (rather than competitive nature of interactions) [12]. Other studies investigated threats to security and trustworthiness resulting from malicious attacks (i.e. vandalism)[22] and the organizational mechanisms used by Wikipedia to combat such attacks [26]; yet these studies do not consider threats resulting from benevolent contributors. A relevant strand of the literature has looked at conflicts of opinions between contributors [13], [12]. However, the focus is on the result of these conflicts on content quality rather than the competitive mechanisms driving them. In summary, while peer-production projects, and in particular Wikipedia, have attracted sig-

nificant attention within the research community, to the best of our knowledge, the competitive dynamics have not been investigated.

In order to better understand collaboration in Wikipedia and capture the competitive nature of interactions, we turn to game theory. Our underlying assumption is that a contributor's goal is to maximize ownership of content sections, such that content "owned" (i.e. originated) by that user survived the most recent revision of the page. We model contributors' interactions to increase their content ownership, as a non-cooperative game. A contributor's motivation for trying to maximize her ownership within a certain topical page could be based on the need to express one's views or to increase her status in the community; and competition could be the result of battles between opposing viewpoints (e.g. vandals and those seeking to ensure trustworthiness of content) or consequences of power struggles. The utility of each contributor in the non-cooperative game is the ownership in the page, defined as the fraction of content owned by the contributor in the page. Each contributor suffers a cost of contribution which is the effort expended towards making the contribution. The objective is then to determine the optimal strategies, i. e., the optimal number of contributions made by each contributor, so that her *net utility* is maximized. Here, the net utility is the difference between the utility (a measure of the ownership) and the cost (a measure of the effort expended). The optimal strategies are determined by determining the Nash equilibrium of the non-cooperative game that models the interactions between the contributors. We determine the conditions under which the Nash equilibrium of the game can be achieved and find its implications on the contributors' expertise levels on the topic. We report of an empirical analysis of Wikipedia that validates the model's predictions. The key results brought forth by our analysis include

- The ownership of contributors increases with the decreasing levels of effort expended by the contributor on the topic.
- Contributors expending equal amount of effort end up with equal ownership.

The rest of the paper is organized as follows. The non-cooperative game that models the interactions between contributors is described in Section 2. We then use Wikipedia data to validate the modeling in Section 3. We then discuss in Section 4 the relevance of our analysis and modeling to trust worthy collaboration and vandalism. Conclusions are drawn in Section 5 along with pointers to future directions.

2 User Contribution as a Non-Cooperative Game

We model the interactions of the N content contributors to a page (i.e., users) as a non-cooperative game. The strategy set for each contributor is the amount and type of contribution she makes. Table 1 describes the notations used in our analysis. and their descriptions.

Let x_i represent the content owned by the ith user in the current version of the page. We define the utility, u_i, as the fraction of content owned by the ith

Table 1. Variables used in the analysis in this paper.

Notation/Variable	Description
N	Number of users or content contributors
x_i	The amount of content owned by the ith user
β_i	Effort expended by user i to make unit contribution
u_i	The fractional ownership held by the ith user
n_i	Net utility of contributor i
1	The all-one vector
I	The identity matrix

contributor, and is given by

$$u_i = \frac{x_i}{\sum_{j=1}^{N} x_j}. \tag{1}$$

The objective of contributor i is to determine the optimal x_i so that u_i is maximum.

It is observed from (1) that the optimal x_i that maximizes u_i is $x_i = \infty$. This is because the utility function is an increasing function of x_i. Intuitively, this result occurs because every time the ith user makes a contribution, his/her ownership increases. However this results in reduction in the ownership of other contributors, to counter which, they attempt to make additional contributions (by increasing their respective x_k's). This, in turn, reduces the ownership of contributor i, thereby causing him/her to further increase x_i to increase ownership. This process continues ad infinitum resulting in $x_i \to \infty$, $\forall$ i. This degenerate scenario can be mitigated as follows.

Suppose the ith contributor expends an effort, β_i, to make a unit contribution. For instance, β_i can be the cost incurred by the ith user, in terms of time and effort spent in learning the topic and in posting content on a Wiki page. Therefore, the ith contributor expends a total effort $\beta_i x_i$, to achieve x_i amount of content ownership in the page. The net utility experienced by the ith contributor, n_i, can be written as the difference between utility of contributor i, given by (1) and the total effort expended by contributor i, i. e.,

$$n_i = u_i - \beta_i x_i = \frac{x_i}{\sum_{j=1}^{N} x_j} - \beta_i x_i. \tag{2}$$

It is observed that the net utility obtained by the ith contributor not only depends on the strategy of the ith contributor (*i.e.*, x_i), but also on the strategies of all the other contributors (*i.e.*, x_j, $j \neq i$). This results in the non-cooperative game of complete information [29] between the contributors. The optimal x_i, $\forall$ i (termed as x_i^*), which is determined by maximizing n_i in (2), is then the Nash equilibrium of the non-cooperative game where no contributor can make a unilateral change.

Applying the first order necessary condition to (2), x_i^* is obtained as the solution to

$$\left.\frac{\partial n_i}{\partial x_i}\right|_{x_i = x_i^*} = \frac{\sum_{\substack{k=1 \\ k \neq i}}^{N} x_k^*}{\left(\sum_{j=1}^{N} x_j^*\right)^2} - \beta_i = 0, \forall i \tag{3}$$

subject to the constraints $x_i^* \geq 0, \forall\, i$. From (3), we obtain $\frac{\partial^2 n_i}{\partial x_i^2} = -\frac{2\sum_{\substack{k=1 \\ k \neq i}}^{N} x_k}{\left(\sum_{j=1}^{N} x_j\right)^3} <$ $0, \forall\, i$, when $x_i \geq 0$. Thus, n_i is a concave function of x_i and x_i^*, which solves (3) subject to $x_i^* \geq 0, \forall\, i$, is a local as well as a global maximum point. In other words, according to [28], *the non-cooperative game has a unique Nash equilibrium*, $\mathbf{x}^* = \begin{bmatrix} x_1^* & x_2^* & \cdots & x_N^* \end{bmatrix}^T$, obtained by numerically solving the system of N non-linear equations specified by (3). However, to study the effect of the effort levels (β_i's) on the strategies of the contributors, it is desirable to obtain an expression that relates the vectors, $\mathbf{x}^*$, $\mathbf{x} = [x_i]_{1 \leq i \leq N}$ and $\boldsymbol{\beta} = [\beta_i]_{1 \leq i \leq N}$.

Re-writing (3),

$$\left(\sum_{j=1}^{N} x_j^*\right)^2 - \alpha_i \sum_{\substack{j=1 \\ j \neq i}}^{N} x_j^* = 0, \forall N, \tag{4}$$

where $\alpha_i \triangleq 1/\beta_i$. Eqn. (4) can be written as

$$(\mathbf{x}^*)^T \mathbf{1}\mathbf{1}^T \mathbf{x}^* \mathbf{1} - \mathbf{D}_\alpha \left(\mathbf{1}\mathbf{1}^T - \mathbf{I}\right) \mathbf{x}^* = \mathbf{0}, \tag{5}$$

where $(.)^T$ represents the transpose of a vector or a matrix, $\mathbf{D}_\alpha$ is the diagonal matrix $\mathbf{diag}\,(\alpha_1, \alpha_2, \cdots, \alpha_N)$, $\mathbf{1}$ is the column vector in which all entries are one, $\mathbf{0}$ is the column vector in which all entries are zero and $\mathbf{I}$ is the identity matrix.

It can be easily verified the vectors, $\mathbf{y}_1 = \begin{bmatrix} \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \frac{1}{\sqrt{N}} & \cdots & \frac{1}{\sqrt{N}} \end{bmatrix}^T$ and for $j = 2, 3, \cdots, N$, $\mathbf{y}_j = \begin{bmatrix} y_{1j} & y_{2j} & y_{3j} & \cdots & y_{(N-1)j} & y_{Nj} \end{bmatrix}^T$, where

$$y_{kj} = \begin{cases} \frac{1}{\sqrt{j(j-1)}} & k < j \\ -\frac{j-1}{\sqrt{j(j-1)}} & k = j \\ 0 & k > j, \end{cases} \tag{6}$$

form a set of orthonormal eigen vectors to the matrix, $\mathbf{1}\mathbf{1}^T$. The eigen value corresponding to $\mathbf{y}_1$ is N and those corresponding to $\mathbf{y}_2, \cdots, \mathbf{y}_N$ are 0s. Let $\mathbf{P} = [\mathbf{y}_1 | \mathbf{y}_2 | \cdots | \mathbf{y}_N]$. Then, $\mathbf{P}$ is an orthogonal matrix and by orthogonality transformation,

$$\mathbf{P}^T \mathbf{1}\mathbf{1}^T \mathbf{P} = \mathbf{D} = \mathrm{diag}\,(N, 0, 0, \cdots, 0). \tag{7}$$

Let $\mathbf{z} = \begin{bmatrix} z_1 & z_2 & z_3 & \cdots & z_{N-1} & z_N \end{bmatrix}^T$. Since the eigen vectors of a matrix form a basis for the $N-$dimensional sub-space [27], the vector, $\mathbf{x}^*$, can be written as $\mathbf{x}^* = \mathbf{P}\mathbf{z}$. A similar expression has been solved in [10] in the context of pricing in wireless networks and we outline here the key steps to determine the optimal $\mathbf{x}^*$.

– Using $\mathbf{x}^* = \mathbf{Pz}$ in (5) and (7), we obtain

$$\mathbf{z}^T \mathbf{Dz1} - \mathbf{D}_\alpha \left(\mathbf{11}^T - \mathbf{I}\right)\mathbf{Pz} = \mathbf{0}. \tag{8}$$

– The above is a set of non-linear equations in $\mathbf{z}$, in which the kth equation depends on z_1 and z_j, $k \leq j \leq N$. Solving the non-linear equations by backward substitution [27], z_k, $2 \leq k \leq N$ can be written in terms of z_1 as

$$\frac{z_k}{\sqrt{k(k-1)}} = \frac{Nz_1^2}{k(k-1)}\left[\frac{k}{\alpha_k} + \sum_{j=k+1}^{N}\frac{1}{\alpha_j}\right] - \frac{z_1}{\sqrt{N}}\frac{N(N-1)}{k(k-1)}. \tag{9}$$

– Using (9) to replace all z_k's in terms of z_1 in the set of non-linear equations in (8), z_1 can be obtained as

$$z_1 = \frac{N-1}{\sqrt{N}}\frac{1}{G}, \tag{10}$$

where

$$G \triangleq \sum_{j=1}^{N}\frac{1}{\alpha_j}. \tag{11}$$

– Combining (9) and (10),

$$\frac{z_k}{\sqrt{k(k-1)}} = \frac{(N-1)^2}{k(k-1)}G^{-1}\left[G^{-1}\left(\frac{k}{\alpha_k} + \sum_{j=k+1}^{N}\frac{1}{\alpha_j}\right) - 1\right] \; 2 \leq k \leq N. \tag{12}$$

– Using the facts $\mathbf{x}^* = \mathbf{Pz}$, and $\alpha_i = \frac{1}{\beta_i}$ in (10) and (12), the unique Nash equilibrium of the non-cooperative game can be obtained as

$$x_i^* = \frac{\sum_{j=1}^{N}\beta_j - (N-1)\beta_i}{\left(\sum_{j=1}^{N}\beta_j\right)^2}. \tag{13}$$

Note that the unique Nash equilibrium $\mathbf{x}^*$, is feasible, *i.e.*, $x_i^* > 0$, $\forall\ i$ if and only if

$$(N-1)\beta_i < \sum_{j=1}^{N}\beta_j. \tag{14}$$

The utility (ownership) of contributor i at the Nash equilibrium, u_i^*, can then be obtained from (1) and (13) as,

$$u_i^* = \left[1 - \left(\frac{(N-1)\beta_i}{\sum_{j=1}^{N}\beta_j}\right)\right]^+, \tag{15}$$

where $x^+ = \max(x, 0)$. It is observed that the ownership u_i^* is non-zero if and only if (14) is satisfied, i.e., if the Nash equilibrium is feasible. The condition in (14) and the expression in (15) have the following interesting implications.

34

– From (15), the ownership of contributors depend on the β_j of *all the contributors*. This is intuitively correct in a peer production project like Wikipedia because contributions are made by multiple users and the ownership held by a user will depend on the effort of all the users that worked together in making contributions to the page.

– The expression in (15) indicates that contributors who expend smaller effort have larger ownership and those who expend larger effort have low ownership, i.e., the fractional content ownership is a decreasing function of the effort expended.

– Asymptotically, i.e., as the number of contributors, N, becomes large, the ownership, u_i^* in (15), can be written as

$$u_i* = \left(1 - \frac{\beta_i}{E[\boldsymbol{\beta}]} \right)^+ , \tag{16}$$

where $E[\boldsymbol{\beta}] \triangleq \frac{1}{N} \sum_{j=1}^{N} \beta_j$, is the *average effort* of all the users that make contributions to the page. From (16), only those contributors for whom $\beta_i < E[\boldsymbol{\beta}]$, i.e., only those contributors whose effort is below the average effort expended in posting content to a page, end up with non-zero ownership. In other words, given the effort involved in making a contribution, and the ease in which others can overwrite one's contributions, only those who expend less effort in making their contributions than the average effort required, end up with non-zero ownership.

3   Empirical Validation with Data

While the non-cooperative game theoretic models developed in Section 2 are based on intuitive notions of ownership and effort, it is necessary to validate these with real data from contributions to Wikipedia articles. We require a set of Wikipedia articles with data on: (a) the content "owned" by contributors at each revision (which can be analogous to the utility, u_i^* in (15) and (b) the cumulative effort exerted by each contributor (including all of his/her contributions) up to each revision, which can represent the effort, β_i, used in the expressions in (12) and (15). We use the data set from Arazy *et al* [13], who explored automated techniques for estimated Wikipedia contributors' relative contributions. The data set in [13] includes nine articles randomly selected from English Wikipedia. Each article was created over an average period of 3.5 years. Section 3.1 presents the details of the data set in [13] and Section 3.2 provides a validation of the same against the models developed in Section 2.

3.1   Extracting data from Wikipedia Articles

The content "owned" by contributors at the end date of each article period was calculated using the method described in [13]. A sentence was employed as the unit of analysis, and each full sentence was initially owned by the contributor who added it. As content on a wiki page evolves, a contributor may

Table 2. The list of articles for the data set in [13] and their attributes.

Article title	Start Date (MM/DD/YYYY)	End Date (MM/DD/YYYY)	Duration (years)	Edits	Unique Editors
Aikodo [1]	11/29/2001	06/13/2004	2.5	72	62
Angel [2]	11/30/2001	12/09/2005	4.0	341	277
Baryon [3]	02/25/2002	08/25/2005	3.5	73	62
Board Game [4]	11/04/2001	12/30/2004	3.2	220	155
Buckminster Fuller [5]	12/13/2001	07/14/2004	2.6	65	55
Center for Disease Control and Prevention [6]	10/16/2001	03/05/2006	4.4	65	58
Classical Mechanics [7]	06/06/2002	08/13/2006	4.2	202	165
Dartmouth College [8]	10/01/2001	08/26/2004	2.9	70	55
Erin Brockovich [9]	09/24/2001	02/02/2006	4.4	59	54
Total			31.7	1167	943
Average			3.5	129.7	104.8

lose a sentence when more than 50% of that sentence was deleted or revised. A contributor making a major revision to a sentence can take ownership of that sentence. The algorithm tracks the evolution of content, recording the number of sentences owned by each contributor at each revision, until the study's end date. The effort exerted by each contributor was, too, based on the method and data set described in [13]. Two research assistants worked independently to analyze every "edit" made to the 9 articles in the sample set and record: contributor's ID; the type of each "edit" to the wiki page (the categories used included: add new content, improve navigation, delete content, proofread, and add hyperlink); the scope of each edit (on a 5-point scale, from minor to major). For example, a particular edit might be categorized as major addition of new content. The two assessors reviewed the "History" section of articles (where Wikipedia keeps a log of all changes to a page), comparing subsequent versions. Once the assessors completed their independent work, and inter-rater agreement levels were calculated (yielding very high levels of agreement), the average of the two assessors was used in the analysis. Finally, the above data set was used to obtain the following parameters on each Wikipedia article listed in Table 2:

- The number of exclusive contributors/users (N)
- The total effort expended by the ith user $(1 \le i \le N)$, s_i
- The number of edits made by the ith user $(1 \le i \le N)$, e_i
- The number of sentences owned by the ith user $(1 \le i \le N)$, p_i.

The following subsection provides a detailed explanation on how we use these parameters to verify the game theoretic analysis described in Section 2.

3.2 Numerical verification of the analysis

Using the set of parameters obtained from the pages in Table 2, listed in Section 3.1, we compute the effort expended by user i for unit contribution, β_i, as

$$\beta_i = \frac{s_i}{e_i}. \tag{17}$$

Using the β_i's thus obtained, we use the expression in (15) to determine the estimated fractional ownership on the Wikipedia page, that will be held by each contributor. We compare this with the fraction $\frac{p_i}{\sum_{j=1}^{N} p_j}$ Figs. 1, 2 and 3 show
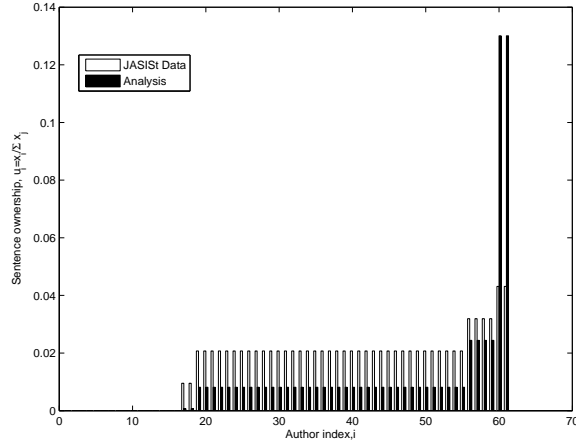
Fig. 1. Ownership of contributors obtained by the game theoretic analysis presented in Section 2 (described by the legend, "Analysis") and the data obtained from Wikipedia pages according to the algorithm in [13] (described by the legend, "JASIST data"), for the page, "Aikido". Contributors/Authors are indexed according to the decreasing order of effort, β_i's.

the comparison between the ownership obtained according to the game theoretic analysis described in Section 2 and that given by the data set in [13] for the Wikipedia pages, "Aikido", "Board Game" and "Erin Brockovich", respectively. For this first analysis, we anonymized the data set, indexing the users in the decreasing order of β_is. We find that the patterns in the empirical data and that of the game-theoretic model closely match one another. In particular, the empirical data validates the following predictions made by the game theoretic model in Section 2[1].

---

[1] These trends were observed not only for the three articles shown in Figs. 1-3 but also for all the nine articles listed in Table 2. We show results for three articles here due to lack of space.
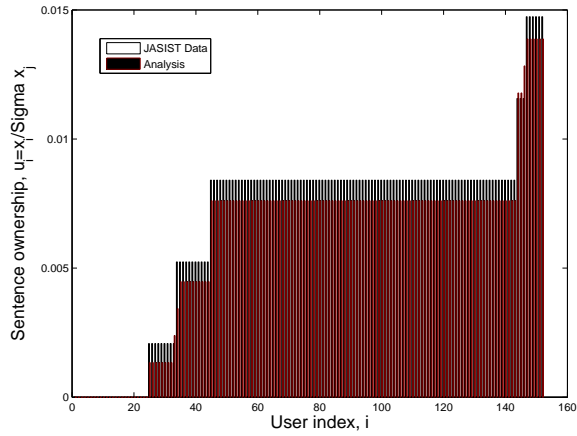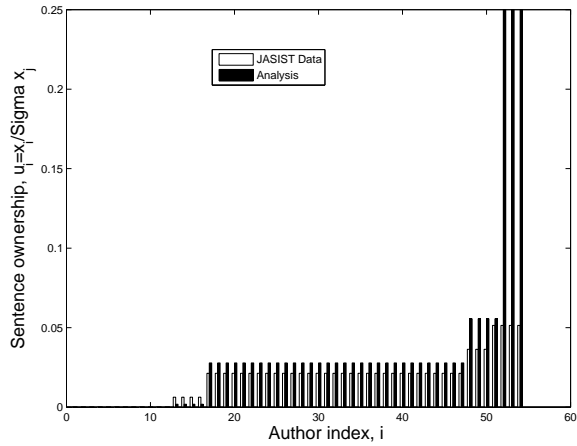
Fig. 2. Ownership of contributors obtained by the game theoretic analysis presented in Section 2 (described by the legend, "Analysis") and the data obtained from Wikipedia pages according to the algorithm in [13] (described by the legend, "JASIST data"), for the page, "Board Game". Contributors/Authors are indexed according to the decreasing order of effort, β_i's.

Fig. 3. Ownership of contributors obtained by the game theoretic analysis presented in Section 2 (described by the legend, "Analysis") and the data obtained from Wikipedia pages according to the algorithm in [13] (described by the legend, "JASIST data"), for the page, "Erin Brockovich". Contributors are indexed according to the decreasing order of effort, β_i's.

38

1. **Equivalence classes:**
   (a) Let the users be classified into equivalence classes according to their fractional ownership, i.e., all users having equal fractional ownership in the Wikipedia page belongs to the same equivalence class. It is observed that each page has five to six equivalence classes. For instance, Aikido, has five equivalence classes (Fig. 1) and Board game (Fig. 2) and Erin Brockovich (Fig. 3), have six equivalence classes each. *Note that the number of equivalence classes obtained from the data is the same as that predicted by the game theoretic analysis described in Section 2.*
   (b) From (15), $u_i^* = u_j^*$ if and only if $\beta_i^* = \beta_j^*$. This indicates that the distribution of the data into number of equivalence classes applies not only to fractional ownership, but also to the effort expended by users. In other words each Wiki page is expected to have five to six categories of contributors/users. A more detailed analysis of the distribution suggests that the majority of users fall into the equivalence middle classes, while the classes on the extreme representing very low and very high levels of effort (and content ownership) comprise of relatively few users. *While the above can be inferred from the data alone, the game theoretic analysis provides a mathematical framework that validates this observation.*
2. **Non-zero ownership:** It is observed from (15) that $u_i^* = 0$ if and only if the condition in (14) is violated. The number of users in our sample data with zero ownership matches the number predicted by the game-theoretic model thus providing validation for the condition (14) (at least for the Wikipedia pages included in our analysis). *Again, it is observed that the relation between the number of users with zero ownership and their corresponding β_i's could have been inferred from the data alone, the game theoretic analysis presented in Section 2 provides a mathematical framework to model this phenomenon.*

After establishing that the general trend (i.e. anonymized data) for the empirical data and the model's predictions match one another, we perform a more detailed analysis where we pay attention to users' identities. That is, we organize both data sets, namely the fractional ownership data taken directly from [13] and the ownership values our model in Section 2 predicted, for each user. We then calculate the correlation between the two data sets, using the Pearson's correlation coefficient [23]. The result of the analysis for the nine articles in our data set is presented in Fig. 4. As could be seen from the figure, correlation coefficients range between 0.47 and 0.88, representing moderate-high correlation. When combining the entire data from the nine articles into a single data set, the Pearson correlation was 0.65 (with a $p-$value, $p \approx 0.04$). Therefore, we now proceed to verify if the discrepancies in the values of the ownership obtained by the game theoretic analysis and that obtained from the data can be offset by establishing a linear fit that maps the set of values obtained by analysis to the ones obtained from the data.

Let $\mathbf{a} \triangleq \begin{bmatrix} a_1 \ a_2 \ a_3 \cdots a_N \end{bmatrix}$ represent the ownership of the contributors obtained by the game theoretic analysis and let $\mathbf{d} \triangleq \begin{bmatrix} d_1 \ d_2 \ d_3 \cdots d_N \end{bmatrix}$ represent the ownership of the contributors obtained from the data as described in [13].
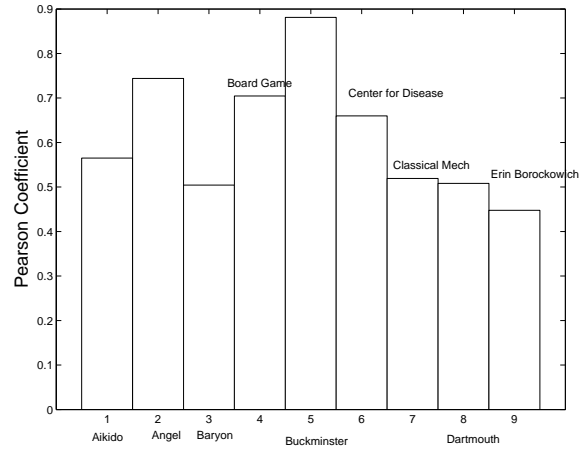
Fig. 4. Pearson correlation co-efficient between the values of the fractional ownership, u_i^*, obtained from the data in [13] and that obtained by the analysis in Section 2
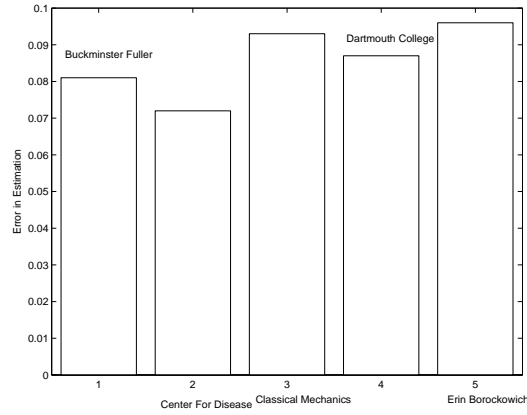

Fig. 5. Estimation error for the linear fit by the method of least squares. The values of the fractional ownership, u_i^*, obtained for the pages, "Aikido", "Angel", "Baryon" and "Board Game" are used as training data for the linear fit.

For each page, we fit a function

$$\hat{d}_i = \rho a_i + \delta \ 1 \le i \le N, \tag{18}$$

where the parameters ρ and δ are obtained by the method of least squares [27]. We use the values for the pages "Aikido", "Angel", "Baryon" and "Board Game" as the training data to obtain ρ and δ. We then use the values of ρ and δ thus obtained to determine $\hat{d}_i$ for the other five pages. We then compare $\hat{d}_i$ and d_i and compute the estimation error for each page. Fig. 5 shows the estimation error for the remaining five pages. It is observed that the error is between 7-9%. The error for the training set of data was found to be around 5%. *This indicates that the game theoretic analysis presented in Section 2 models the contributor interactions in peer production projects like Wikipedia accurately upto a linear scaling factor.*

4   Trustworthy Collaboration and Vandalism

An important insight provided by our non-cooperative game model (and validated by our empirical analysis) is that only contributors with below-average effort levels are able to maintain fractional ownership on wiki pages. That is, by and large only the edits made by contributors who exert little effort survive the collaborative authoring process. In Section 1, we referred to two key concerns that are associated with trustworthy collaboration in peer-production projects: (a) a risk that non-experts will contribute content of low quality, and (b) a threat that malicious participants would vandalize Wikipedia pages. In spite of these serious concerns, the content on Wikipedia articles is generally of high quality and Wikipedia maintains the status as one of the most reliable sources of information on the web [20]. How then, does Wikipedia maintain high-quality content in the face of threats of low-quality or malicious contributions? Our results can have important implications for investigation of trustworthy collaboration on Wikipedia (and more broadly, in peer-production projects). In the sections that follow, we provide two interpretations of our results that help explain how the threats highlighted above are mitigated.

1. **Trustworthiness/Quality of Wikipedia pages:** The first interpretation of the model and its empirical validation involves the concern of non-expert, low quality contributions eroding the trustworthiness of peer-produced product. This interpretation suggests that low effort is associated with greater likelihood of content survival due to a skill advantage: contributors who are experts in their field of contribution expend less effort, and their contributions are of higher quality [11]. Thus, the effort associated with contribution is inversely related with its quality and consequently with its likelihood of survival of subsequent editing.
2. **Vandalism:** The second interpretation concerns the danger of vandalism activities reducing the trustworthiness of the peer-produced products. Since the underlying Wiki mechanisms allow any editor to easily revert the edits of

other contributors, the effort involved in vandalistic edits is higher than the effort of reverting such edits. Thus, high effort is associated with vandalism and relatively lower effort is linked to correction of vandalism. The game theoretic analysis presented in Section 2 predicts that the contributions made by users expending large effort do not survive the edit process and end up with zero ownership. Therefore, most vandalistic edits would not survive over time, as also observed in [21], [33], [32].

In summary, following on the intuition observed in [21], [33], [32], we modeled competition between players as a non-cooperative game, where a player's utility is associated with surviving fractional content owned, and cost is a function of effort exerted. Broader design implications emerging from this interpretation include the need to make version control mechanisms not only highly usable, but also highly open and egalitarian, and accessible to participants in a peer-production process. In addition, these insights suggest the importance of concurrent use of other quality control mechanisms, including user-designated alerts (where users are notified when changes are made to an article, or other part of the collaboratively-created product); watch lists (where users can track certain articles); and IP or user blocking in cases where repeated attacks from the same source are deemed to be acts of vandalism. The combination of these mechanisms make three important contributions to the trustworthiness of peer-production projects: first, their existence deter potential vandals; second, they reduce the costs of identifying and responding quickly to attacks; and third, they enable users to easily revert the consequences of vandalism .

5 Conclusion

To better understand the success of peer production, we developed a non-cooperative game theoretic model of the creation of Wikipedia articles. The utility of a contributor was her relative ownership of the peer-produced product that survived a large number of iterations of collaborative editing. The work presented here contributes to better understanding of the trustworthiness of peer-production by

- Solving the game and demonstrating the conditions under which a Nash equilibrium exists, showing that asymptotically only users with below average effort would maintain ownership
- Empirically validating the model, demonstrating that only users with below average effort would maintain ownership, as well as showing editors' equivalence classes
- Offering interpretations and implications for research on trustworthy peer-production (in terms of expertise and vandalism).

To the best of our knowledge, this is the first modeling of user interactions on Wikipedia as a non-cooperative game. Our analysis points to the benefits of deploying multiple mechanisms which afford the combination of large-scale and

low-effort quality control as way to ensure the trustworthiness of products created through web-based peer-production. Further research is needed to analyze the effectiveness of each of these mechanisms, and to address other aspects of peer production through game theoretic analysis.

Acknowledgment

References

1. http://en.wikipedia.org/wiki/Aikido
2. http://en.wikipedia.org/wiki/Angel
3. http://en.wikipedia.org/wiki/Baryon
4. http://en.wikipedia.org/wiki/Board_Game
5. http://en.wikipedia.org/wiki/Buckminster_Fuller
6. http://en.wikipedia.org/wiki/Centers_for_Disease_Control_and_Prevention
7. http://en.wikipedia.org/wiki/Classical_Mechanics
8. http://en.wikipedia.org/wiki/Dartmouth_College
9. http://en.wikipedia.org/wiki/Erin_Borockowich
10. Anand, S., Sengupta, S., Chandramouli, R.: Price bandwidth dynamics for WSPs in heterogeneous wireless networks. Technical Report, Stevens Institute of Technology (2011)
11. Anthony, D., et al: The quality of open source production:Zealots and good samaritans in the case of Wikipedia. Rationality and Society (2007)
12. Arazy, O., Nov, O., Patterson, R., Yeo, L.: Information quality in Wikipedia: The effects of group composition and task conflict. Journal of Management Information Systems 27, 7198 (2011)
13. Arazy, O., Stroulia, E., Ruecker, S., Arias, C., Fieurontino, C., Ganev, V., Yao, T.: Recongnizing contributions in Wikis: Authorship categories, algorithms and visualizations. Journal Of American Society for Information Science and Technology 61(6), 1166–1179 (Jun 2010)
14. Benkler, Y.: The Wealth of Networks: How Social Production Transforms Markets and Freedom. Yale University Press (2006)
15. C. Wagner, C., Majchrzak, A.: Enabling customer-centricity using Wikis and the Wiki way. Journal of Management Information Systems 23, 17–43 (2007)
16. Eisenhardt, K.M., Martin, J.A.: Dynamic capabilities: What are they? Strategic Management Journal 21, 1105–1121 (2000)
17. Forte, A., Park, T.: How people assess cooperatively authored information resources. Proceedings, Eighth Annual International Symposium on Wikis and Open Collaboration (2012)
18. Grewal, R., Lilien, G.L., Mallapragada, G.: Location, location, location: How network embeddedness affects project success in open source systems. Management Science 52, 1043–1056 (2006)
19. Hippel, E.V., Krogh, G.V.: Open source software and the "private collective" innovation model: Issues for organization science. Organization Science 14, 209–223 (2003)

20. Kittur, A., et al: Can you ever trust a Wiki?: Impacting perceived trustworthiness in Wikipedia. Proceedings, ACM Conference on Computer Supported Cooperative Work (2008)
21. Kittur, A., Suh, B., Pendleton, B.A., Chi, E.H.: He says, she says: Conflict and coordination in Wikipedia. Proceedings, ACM SIGCHI conference on Human factors in computing systems pp. 453–462 (April 2007)
22. Krogh, G.V., Hippel, E.V.: The promise of research on open source software. Management Science 52, 975–983 (2006)
23. Lakhani, K., Wolf, R.: Why hackers do what they do: Understanding motivation effort? Perspectives in Free and Open-Source Software, B. F. J. Feller, S. Hissam, & K. Lakhani Ed., ed: MIT Press (2005)
24. Lee, G.K., Cole, R.E.: From a firm-based to a community-based model of knowledge creation: The case of the Linux Kernel development. Organization Science 14, 633–649 (2003)
25. Leuf, B., Cunningham, W.: The Wiki way: Quick Collaboration on the Web: Reading, Massachusetts. Addison-Wesley (2001)
26. Majchrzak, A.: Where is the theory in Wikis? MIS Quarterly 33, 18–20 (2009)
27. Meyer, C.D.: Matrix Theory and Applied Linear Algebra. SIAM (1972)
28. Nash, J.: Equilibrium points in N-person games. National Academy of Sciences 36, 48–49 (1950)
29. Neumann, J.V., Morgenstern, O.: Theory of Games and Economic Behavior. Princeton University Press (1944)
30. Nov, O., Kuk, G.: Open source content contributors' response to free-riding: The effect of personality and context. Computers in Human Behavior 24(6), 2848–2861 (2008)
31. Peng, D.X., Schroeder, R.G., Shah, R.: Linking routines to operations capabilities: A new perspective. Journal of Operations Management 26, 730–748 (2008)
32. Stvilia, B., Twidale, M.B., Smith, L., Gasser, L.: Information quality work organization in Wikipedia. Journal of the American Society for Information Science and Technology 59(6), 983–1001 (June 2008)
33. Suh, B., Chi, E.H., Kittur, B.A.P..A.: Us vs them: Understanding social dynamics in Wikipedia with revert graph visualizations. IEEE Symposium on Visual Analytics Science and Technology (VAST'2007) pp. 163–170 (October 2007)
34. Teece, D.J., Pisano, G., Shuen, A.: Dynamic capabilities and strategic management. Strategic Management Journal 18, 509–533 (1997)
35. Towne, W.B., et al.: Your process is showing: Controversy management and perceived quality in Wikipedia. Proceedings, ACM Conference on Computer Supported Cooperative Work (2013)
36. Wagner, C.: Wiki: A technology for conversational knowledge management and group collaboration. Communications of the Association for Information Systems 13, 265–289 (2004)
37. Winter, S.G.: Understanding dynamic capabilities. Strategic Management Journal 24, 991–995 (2003)