Social Rewarding in Wiki Systems – Motivating the Community

Bernhard Hoisl[1], Wolfgang Aigner[2], Silvia Miksch[2]

[1] Institute of Software Technology and Interactive Systems, Vienna University of Technology, Favoritenstr. 9-11/188, 1040 Vienna, Austria
e0252748@student.tuwien.ac.at

[2] Department of Information and Knowledge Engineering, Danube University Krems, Dr.-Karl-Dorrek-Str. 30, 3500 Krems, Austria
{wolfgang.aigner, silvia.miksch}@donau-uni.ac.at

Abstract. Online communities have something in common: their success rise and fall with the participation rate of active users. In this paper we focus on social rewarding mechanisms that generate benefits for users in order to achieve a higher contribution rate in a wiki system. In an online community, social rewarding is in the majority of cases based on accentuation of the most active members. As money cannot be used as a motivating factor others like status, power, acceptance, and glory have to be employed. We explain different social rewarding mechanisms which aim to meet these needs of users. Furthermore, we implemented a number of methods within the MediaWiki system, where social rewarding criteria are satisfied by generating a ranking of most active members.

Keywords: Social Rewarding, Wiki, Online Communities, Motivation, Participation, Contribution.

1 Introduction

Wikipedia – the most famous free encyclopaedia – has grown to the biggest wiki community site where hundreds of thousands of users all around the world post and edit articles in many different languages. The tremendous contribution rate on Wikipedia has led to many problems, like wrong information, copyright violations, or users' misbehaviour, for example, spammers or trolls [16]. Other online communities beside Wikipedia have massive troubles motivating users to participate actively. We are going to present techniques where the fundamental problem of both – reaching a critical mass of active users – are addressed.

On the one hand, Wikipedia has the problem that published information is not checked for its accuracy and legality by a formal process of reviewing. There has to be a large and heavily involved community which is cross-checking and proofing information for its correctness voluntarily. However, the operators of Wikipedia have not only a social but also a legal responsibility to publish only correct and faultless

information to assure their creditability. On the other hand, many online communities have troubles motivating enough users to build an active community. Participation of members is the key factor for a successful online community, and that is why good motivating factors are essential.

As information provided over the Internet is treated like public goods, problems like free riding[1] or social loafing[2] arise. In Wikipedia users are not charged in proportion to their use, therefore it appears rational for people to view articles without contributing anything on their own. If we assume an economic point of view it can be said that a user has costs by publishing an article to Wikipedia (e.g., information acquisition and presentation costs or Internet connection costs) and therefore she/he wants something in return. Extending the benefit for a user so that it exceeds her/his costs is a good starting point to increase participation. With this contribution we are going to focus on an approach to motivate users to participate actively in an online community by making use of a number of different social rewarding techniques [8].

To classify our approach, we will give an overview of related work in the next section. Section 3 will explain the developed social rewarding techniques while section 4 gives an insight on the calculation process of these methods. Section 5 covers the visual appearance of the authors' ranking and the implementation is summarized in section 6. A conclusion is drawn in section 7 containing an outlook on future work.

2 Related Work

There are numerous books and articles about the wiki phenomenon (e.g., [5, 11, 14, 18]). However, most work focuses on technical details, like installing and running a wiki or the revisioning system and its vantages for collaborative information development. Unfortunately, too little attention is paid to investigate users' behaviour in online communities. Some research is done to explain the problem of free riding [1, 6] which is likely to occur in times of the Internet and shared information platforms. There are also studies about communication activities of users in virtual communities [17], but the focus is not on motivational factors for users of online communities.

Which factors are motivating for a human being, was already discussed by Abraham Maslow and his hierarchy of human needs theory [13]. In an article about

---

[1] In this case, free riding means that a user shoulders less than a fair share of the costs of the whole information production of a wiki [3]. If everybody contributes the same value of information to a wiki, nobody free rides. One of the biggest problems is that the value of an information resource to an individual is very subjective and hard to determine.

[2] Social loafing is the phenomenon that persons make less effort to achieve a goal when they work in a group than when they work alone [9]. As the least articles in Wikipedia (like in nearly every other wiki) are written by only one user but in a team the problem of social loafing is likely to occur. The answer to social loafing are motivational factors which are partly solved in the MediaWiki software by the possibility to see which sections of an article belongs to which author. So, a contribution is linked to an author's name and can therefore be evaluated.

using social psychology to motivate contributions to online communities [10] an experiment took place where the problems of under-contribution and social loafing were addressed. In the article, as predicted by theory, individuals contributed when they were reminded of their uniqueness and when they were given specific and challenging goals. As other predictions were disconfirmed, results of the experiment have to be interpreted carefully. An article from the same co-authors [12] focusing on a related topic, tried to manipulate two factors to increase participation in online communities: *similarity* – how similar group members' contributions were and *uniqueness* – how unique members' contributions were within the group. As a result both factors positively influenced participation.

Our approach to increase users' participation in a wiki is based on accentuation and reputation [15]. By motivating many users we want to increase the community so that cross-checking takes place and false information is automatically sorted out. That such an approach of member-maintained communities increases the quantity and quality of contributions was affirmed [4] and empirically tested on Wikipedia [2].

3 Social Rewarding Techniques

In this paper we present social rewarding mechanisms that generate benefits for the users in order to achieve a higher contribution rate in a wiki community. In our case, social reward refers to something that causes a behaviour to increase in intensity. In an online community, social rewarding is in the majority of cases based on accentuation of the most active members. As money cannot be used as a motivating factor, others like status, power, acceptance, and glory have to be employed. We explain different social rewarding methods which aim to meet these needs of users.

The techniques presented are focussing primarily on automatic investigations of quantitative and qualitative characteristics of published articles. As a proof of concept, three social rewarding mechanisms were implemented using the software MediaWiki[3] (which is also used by Wikipedia). Most active members are accentuated by applying these social rewarding methods to calculate a ranking of authors:

- *Amount of References* – This social rewarding method uses Google's SOAP search API to build an index quality number based on three different criteria: the size of a reference, the number of links pointing to this reference and the number of links pointing to the specific article.
- *Rating of Articles* – A user centric evaluation of articles published is still missing in the MediaWiki software. We have implemented an open rating system where users can vote for or against an article (and optionally leave a comment) by making use of a predefined pointing scale.
- *Most Viewed Articles* – Visits of users are counted working with configured parameters.

[3] http://www.mediawiki.org

The two most important criteria for our choice were, on the one hand, to find a good mixture of different methods and, on the other hand, the level of complexity of the implementation process in MediaWiki.

We believe that using a couple of different social rewarding mechanisms will result in better findings for two reasons. The first reason is because data will be retrieved from different sources. Combining these data should result in a better and more plausible result than any other technique alone. The second reason is that many different data sources make it hard for an author to betray. If we only count users' hits it is obvious that authors would try to cheat by visiting their own articles a lot more often than other ones. Of course there have to be control mechanisms, like preventing authors to be counted as visitors of their own articles. But too many restrictions can falsify the *real* behaviour of users which we are trying to measure.

3.1 Amount of References

As in the case of the Wikipedia encyclopaedia the value of an article grows with the amount and quality of used references. An approach to an automated quality check of Internet resources was realized by the help of one of the world's largest search engines: Google. Google can help to detect the quality of an article by figuring out how much sites are linked to a cited reference[4]. If many sites are linking to an Internet resource cited as a reference and this links themselves have a high number on links to them, then information displayed on this site must have at least a basic level of plausibility (this concept is the basis for Google's search algorithm named *PageRank* [7]).

Besides the number of links to a reference another criterion – the size of the reference – is used[5]. We assume that a reference with thousands of sub-sites can be more trusted than a home-made personal web-site with only three pages.

A more global idea is counting the links to a wiki article from outside. By using Google we cannot only check references within articles, but also figure out how many sites outside the wiki are pointing to an article. If there are thousands of links to an article it is likely that this article is valuable to many people. The higher the amount of links to an article is, the higher is the frequency of visitors and readers. Many links are also an indicator for good quality of an article.

Our calculation is influenced by these three directly presented criteria: the number of links pointing to a reference, the size of this reference and the number of links pointing to the specific wiki article. Now a quality index number of an article can be generated which can be used for a basic classification of the references as more or less credibly and which can indicate the publicity of the article. So at least an initial quality check of Internet resources can be realized by using Google's PageRank technique. This attempt tries to rank articles not only by means of quantitative characteristics but also qualitative ones.

---

[4] Entering *link:http://www.tuwien.ac.at* as a search term will result in showing all pages linking to this address.

[5] By inserting *site:http://www.tuwien.ac.at* as a search term the number of sub-sites belonging to this address will be returned.

3.2 Rating of Articles

To distinguish good written articles from bad ones, the user has the possibility to vote for or against it. This is done by asking only one simple question with standardized answer alternatives. Answer possibilities could be: *Yes/No* (”Did you like the article?”); *-5 to +5* (“How relevant was the information shown in this article to you?”) or something similar to that. A text field is inserted giving the user the chance to write in what she/he liked/disliked. So the rating points are quantitatively calculated while the author also gets a personal qualitative feedback.

The rating results are inserted in the discussion page of the article. Users and especially the author her-/himself can have a quick overview why users rated the article positively or negatively. As a next step the author could rewrite the article based on the ideas of the users (certainly users can do this also on their own). In the discussion page the author has the possibility to post answers upon users' comments, thus giving her/him the chance for a justification.

Other aspects that have to be considered are the number of minimal votes needed for a representative result and some sort of protection against multiple votes.

3.3 Most Viewed Articles

The idea behind a list of most viewed articles is that when an article is viewed by many people it is either (1) very informative and very well written with good background knowledge of the author, or (2) it has a highly interesting theme for a broad range of people. If we assume case one then it can be said that articles which have a high rate of hits or visits help to achieve a good reputation for their authors.

A list of most viewed articles can be an overall list of most viewed articles ever, separated by a certain amount of time, or they can be categorized by their topic. A list of most viewed articles ever can be a good idea, although there will certainly not be very much fluctuation among the top articles in the list. To avoid this behaviour most viewed articles of the month or week can be a solution.

The following section explains how these directly presented social rewarding methods are combined to find out most active users.

4 Calculation

As said before our developed social rewarding techniques are focussing on accentuation of the most active members in a community. This is done by highlighting the most productive authors in a ranking. In the former chapter we introduced the social rewarding methods which are used for the calculation of such a listing. Now it is time to explain the two-step calculation process.

4.1 Revision Basis

Each of the three social rewarding mechanisms computes points for a single revision[6] of every article. This is done by comparing the value of the specific revision with the average value of all revisions in the wiki (see equation 1).

$$avg_{Rj} = \frac{\sum_{i=1}^{n} r_{ij}}{n}. \quad (1)$$

R is a set of all revisions where r_{ij} is the value of the j social rewarding mechanism of revision i:

$$R_i = \{r_{i1}, r_{i2}, r_{i3,\ldots,} r_{ij}; 1 \leq j \leq 3\}. \quad (2)$$

For example, for the social rewarding technique *Most Viewed Articles* all visits to a revision are counted. Let us assume article *A* in revision *7* has *20* views. The average value of views of all revisions is *30*. So revision *7* of article *A* has only *66.67%* of the overall average views. As we want to credit every revision with a certain amount of points according to their visits, a scale must be predefined to set the intervals. In our scale a value of *66.67%* would be graded with *2* out of *5* points[7].

This example of point assignment is done for every revision and for every social rewarding method. For mechanism *Rating of Articles* users vote for an article by assigning *0* to *5* points. For the technique *Amount of References* the number of links pointing to a reference, the size of this reference and the number of links pointing to the specific article are used as variables. These three criteria are weighted according to users' settings and are compared to a mean value calculated over all revisions.

In the end of the first computation step for every social rewarding method and for every revision points are assigned according to predefined scales. These values are weighted and summed up to an overall value per revision. By looking at equation 3 it can be seen that p_{r_i} are the summed up points for revision i for every social rewarding method j weighted against w_j (which has to be defined in the configuration file).

$$p_{r_i} = \sum_{j=1}^{3} \frac{r_{ij}}{avg_{Rj}} * w_j \,. \quad (3)$$

The allocation of points of revisions to authors is done in the next step.

4.2 Author Basis

As a revision is linked to exactly one author it is now possible to sum up all points of every revision an author has written. This is done by using two methods to weight the result: the length of the edit and the creation time of a revision.

---

[6] Every change made to an article results in a new revision.

[7] These examples use a scale from *0* (worst) to *5* (best) points, but it can be defined as wanted.

A modified set of R is created where r'_{ik} is revision k of article i (equation 4).

$$R'_i = \left\{ r'_{i1}, r'_{i2}, r'_{i3}, \ldots, r'_{ik} \right\}. \tag{4}$$

We assume that the more different a new revision is compared to the former one, the more important were the changes made. It does not matter, if a new revision is extended or shortened – a surplus in content quality is assumed[8]. The difference from one revision to another is counted in bytes. Using equation 5 we get an overall value of size changes from all revisions k from an article i (where $s_{r'ik}$ is the specific size change from one revision to the former one).

$$S_{r'_i} = \sum_{k=1}^{n} \underbrace{abs\left[size\left(r'_{ik}\right) - size\left(r'_{i(k-1)}\right)\right]}_{s_{r'ik}}. \tag{5}$$

The second assumption is that newer revisions count more than older ones. We believe that newer revisions have up-to-date topics and therefore should be weighted higher than ones written long ago. Equation 6 sums up the relative amount of time for all revisions k for an article i ($t_{r'ik}$ is the relative amount of time for one revision).

$$T_{r'_i} = \sum_{k=1}^{n} \underbrace{time\left(r'_{ik}\right) - firstTime\left(r'_i\right)}_{t_{r'ik}}. \tag{6}$$

For all revisions of every article, the differentiation to the former revision, and its age according to the creation date of the article are saved (equation 5 and 6).

Equation 7 defines a subset A of revisions belonging to one author. This means that only revisions from the specific author for whom the calculation takes place are considered. So, for example, $s_{a_{ik}}$ (in equation 8) is the size change from one revision of the author which is divided by the overall size change of all revisions of the article to get a percentage value.

$$A \subseteq R'. \tag{7}$$

In equation 8 for every revision belonging to an author and every criterion (size and time) percentage values are generated which are weighted using a predefined scale (w_S and w_T). Then these two values are multiplied with the specific points calculated in the first step for this revision (p_{r_k}) and both values are summed up. The outcome is a new weighted value for every revision ($p_{a_{ik}}$) which has to be summed up for all articles belonging to an author (p_A).

$$p_A = \sum_{i=1}^{n} \sum_{k=1}^{m} \underbrace{\frac{s_{a_{ik}}}{S_{r'_i}} * w_S * p_{r_k} + \frac{t_{a_{ik}}}{T_{r'_i}} * w_T * p_{r_k}}_{p_{a_{ik}}}. \tag{8}$$

[8] That means we factor out flamers, trolls etc.

This procedure has to be done for all authors, so that in the end every author has one value assigned which is the basis for displaying the ranking. Fig. 1 gives an overview of the two-step calculation process described in this section.

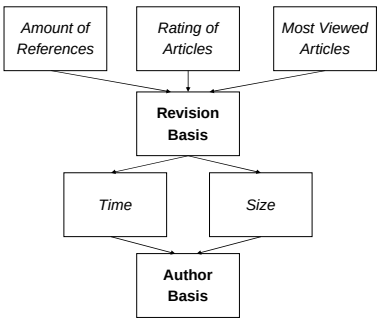

Fig. 1. Two-step calculation process. At first points are computed on a revision basis using the three described social rewarding methods. In a second step the points are weighted according to time and size factors and summed up for an author.

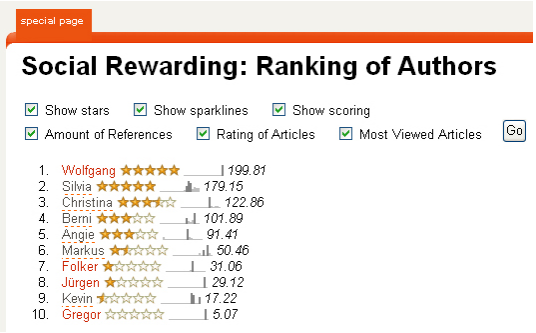

Fig. 2. Screenshot of ranking of authors. Besides the authors' names, stars and sparklines can be seen. The numbers on the right are the achieved scores according to the calculation of the social rewarding methods.

5 Ranking of Authors

For displaying results, various authors' rankings can be generated where the most active one will see her-/himself on the first place (Fig. 2). To support shown results, two well-known data visualization techniques are used: *stars* and *sparklines* [19].

5.1 Stars

Using stars to generate a ranking is well known and an established way to give a quick indication on how good or bad something is. Large Internet sites, like eBay or Amazon and many forum applications use stars as graphical expressions. We recommend using a five star scaling to show the participation rate of a user (displaying half-stars can be activated additionally). As stars are computed on the basis of the participation rate of all other members of the community, they are a good indication for the overall contribution rate of a user.

5.2 Sparklines

Sparklines are *"small, high-resolution graphics embedded in a context of words, numbers, images. Sparklines are data-intense, design-simple, word-sized graphics.*

Sparklines have obvious applications for financial and economic data, by tracking changes over time, showing overall trend as well as local detail" [19].

In this work sparklines are used to show the participation rate of a user over a certain period of time split by predefined intervals. Therefore, the contribution rate is calculated using the three social rewarding mechanisms described earlier. We have chosen sparklines mainly because of their good integration in a context of words and their simplicity. The appearance, intervals, heights, widths, spaces, and colours of the sparklines can be customized by the user.

6 Implementation

We implemented our developed social rewarding techniques as an extension in the MediaWiki system. For setting up the extension a configuration file is used where all variables belonging to our package can be configured (*~100*).

As the computation of the authors' ranking depends strongly on the amount of articles, revisions, and authors it can be very time consuming. Therefore, a caching algorithm was implemented so that the calculation does not have to be done upon every single request. Caching data can either be saved on the file system or in the database. By selecting the latter a history of authors' ranking can be generated.

Most functions of the extension were implemented to be displayed as so-called *SpecialPages*. But also some self-defined markups can be inserted into an article to display information provided by our package. At last, hooks are used for collecting necessary data for the computation process.

7 Conclusion and Future Work

Because under-contribution is a serious problem for many online communities, we have tried in this paper to give an insight on how to motivate users by means of social rewarding techniques. We based our work on the accentuation of most active members in a wiki. To find these users we generated an authors' ranking by making use of calculated points of three developed social rewarding mechanisms: *Amount of References*, *Rating of Articles,* and *Most Viewed Articles*. Several weighting variables influence the ranking. Some of them are configurable; others rely on the quantity, quality, and novelty of the authors' text. Besides the ranking, stars and sparklines are used to visualize the results. As an implementation platform we have chosen MediaWiki in which our social rewarding mechanisms where integrated.

Our approach can be seen as a starting point to develop mechanisms to the important issue of motivating users to participate actively in a wiki system. In no other online community the participation rate of users is more important than in a wiki, because there producers and consumers of the good (namely information) are the same. If too less users produce content and only free ride a wiki community will cannibalize itself. We think that our implementation of social rewarding techniques as a mixture of several methods is a good way to create qualitative high results which are necessary to generate non-monetary incentives for users. Nevertheless, it is a failure

to think that mechanisms we have described in this paper will be sufficient to motivate enough people to form an active community to participate in every wiki. Users must have an intrinsic motivation to contribute to a wiki which with our developed techniques can only be stimulated.

Our project is not publicly released yet and therefore empirical data is not available. For this reason, we are currently planning to evaluate our implemented concepts in a larger setting.

References

1. Adar, Eytan; Huberman, Bernardo A.: Free Riding on Gnutella. Xerox Palo Alto Research Center (2000).
2. Chesney, Thomas: An Empirical Examination of Wikipedia's Credibility. First Monday, volume 11, number 11 (2006).
3. Cornes, Richard; Sandler, Todd: The Theory of Externalities, Public Goods and Club Goods. 2nd ed. Cambridge Univ. Press (1996).
4. Cosley, Dan; Frankowski, Dan; Kiesler, Sara; et. al.: How Oversight Improves Member-Maintained Communities. CommunityLab (2005).
5. Ebersbach, Anna; Glaser, Markus; Heigl, Richard: Wiki: Web Collaboration. Springer, Berlin Heidelberg New York (2005).
6. Feldman, Michal; Papadimitriou, Christos; Chuang, John; et. al.: Free-Riding and Whitewashing in Peer-to-Peer Systems. School of Information Management and Systems, UC Berkeley (2004).
7. Google Technology: http://www.google.com/technology/, retrieved on 2007-02-15.
8. Hoisl, Bernhard: Social Rewarding in Online Communities - A Focus on Wiki Systems. Master thesis, to be published. Vienna University of Technology (2007).
9. Jackson, J. M.; Harkins, S. G.: Equity in Effort: An Explanation of the Social Loafing Effect. In: Journal of Personality and Social Psychology, volume 49, 1199-1206 (1985).
10. Kimberly, Ling; Beenen, Gerard; Ludford, Pamela; et. al.: Using Social Psychology to Motivate Contributions to Online Communities. CommunityLab (2004).
11. Leuf, Bo; Cunningham, Ward: The Wiki Way: Quick Collaboration on the Web. Addison-Wesley Professional (2001).
12. Ludford, Pamela; Cosley, Dan; Frankowski, Dan; et. al.: Think Different: Increasing Online Community Participation Using Uniqueness and Group Dissimilarity. University of Minnesota, Department of Computer Science and Engineering (2004).
13. Maslow, Abraham H.: Motivation and Personality. 3rd ed. HarperCollins Publishers (1987).
14. McFederies, Paul: Technically Speaking: It's A Wiki, Wiki World. In: IEEE Spectrum, volume 43, 88 (2006).
15. Resnick, Paul; Zeckhauser, Richard; Friedman, Eric; et. al.: Reputation Systems: Facilitating Trust in Internet Interactions. Communications of the ACM, volume 43, issue 12, 45-48 (2000).
16. Sanger, Larry: Why Wikipedia Must Jettison Its Anti-Elitism. http://www.kuro5hin.org/story/2004/12/30/142458/25, retrieved on 2007-02-15. Kuro5hin (2004).
17. Schoberth, Thomas; Preece, Jenny; Heinzl, Armin: Online Communities: A Longitude Analysis of Communication Activities. Working Paper in Information Systems (2003).
18. Tapscott, Don; Williams, Anthony D.: Wikinomics: How Mass Collaboration Changes Everything. Portfolio (2006).
19. Tufte, Edward R.: Beautiful Evidence. Graphics Press, 7 – 19 (2006).