

Sadlon, E., Sakamoto, Y., Dever, H. J., Nickerson, J. V. (2008). In Proceedings of the 18th Annual Workshop on Information Technologies and Systems.

## The Karma of Digg: Reciprocity in Online Social Networks

**Abstract:** *What motivates participants to donate their time to collective intelligence web sites? We examine the possible mechanisms that drive the website Digg, an online news consolidator in which users help determine the popularity and visibility of stories and other web-based information. Those who submit stories that become popular also actively read and vote for each other's stories. This behavior can be explained as a form of reciprocity; we show evidence for this and discuss the implications for other communities of practice.*

**Résumé:** *Que motivent les participants à passer du temps sur des sites web communautaires? Nous étudions les mécanismes du site Digg, un agrégateur de nouvelles dans lequel les utilisateurs déterminent la popularité des histoires. Les rédacteurs des histoires qui deviennent populaires lisent et votent pour les histoires des autres. Ce comportement est une forme de réciprocité.*

### 1. Introduction

Web sites now abound that contain information contributed by users: Wikipedia, SourceForge, and a host of others. Some see these sites as reflecting a form of collective intelligence born of a sophisticated and altruistic user base [1]. Others think that, while altruistic behavior may exist, the driving force on many web sites is a transactional reciprocity intended to build reputation [2].

Those who model reciprocity distinguish between its direct form and its indirect form [3-7]. Indirect reciprocity posits that people do good for the group with a general understanding that good will come back: reputation or image systems provide a mechanism for signaling between participants.

Here we analyze the nature of participation on the website Digg. Digg provides a filtering function – stories are posted at the rate of about one a second, and, through a process of user voting (each vote is called a *digg*), stories move into a state of either high popularity or obscurity. Users act as editors, culling stories for this online community. We are interested in what motivates and sustains such a community. Consistent with the theory of distributed cognition [8], we think these complex and useful institutions emerge from a set of readily understandable behaviors.

Our contribution is as follows: we present evidence that there is a small subset of the Digg community that account for all of the stories that become popular. Members of this subset behave differently from the rest of the community, and, in particular, are remarkable in that they balance their time between submitting new stories and voting for the stories that have recently been submitted by others. In other words, they reciprocate with respect to evaluation. We identify mechanisms through which these diggers signal each other. These findings further our understanding of how online communities can quickly grow by using mechanisms that encourage reciprocity, which in turn lead participants to embed themselves in the community, playing the roles of both originators and evaluators.

In the sections that follow, we will provide more background and make several predictions, and then present the results of our data analysis.

### 2. BACKGROUND

Communities in which users volunteer time are sometimes hard to explain with traditional economic theories. In classical economics, labor is traded for payment. In sites such as Digg, labor is contributed with no obvious monetary compensation. Anthropologists, however, have argued that much of culture takes place in non-market settings, and in such settings, reciprocity functions as a motivational force [9].

Furthermore, social psychologists have found that the urge to reciprocate is deeply ingrained [10]. In the context of online communities, it has been shown that rating systems contain evidence of reciprocation [2]. This last work, however, looked at rating systems on commerce sites. Do sites such as Digg, which don't directly sell commercial products, function in the same way?

Just as with many other sites, Digg has an informal reputation system. The number of diggs made by a user, the number of stories submitted, and the percentage of stories submitted that have been promoted to the status of popularity, are all metrics posted on the publicly accessible profile pages of users. Some users acquire celebrity status on the website because they have built up large friendship networks and have acquired an ability to submit stories that will later become popular (see [11] for a more detailed explanation of influence on Digg).

This status is not easy to achieve. Consider that stories are submitted at the rate of about 1 per second. Most users who come to the website only see stories that have already become popular, not stories recently submitted. The recently submitted stories sit on a set of several hundred web pages, and it is unlikely that anyone can easily browse through these pages, or would be motivated to do so. The site has a social networking *friend* construct: once a user becomes the fan of someone, then the user can then see activity of the person they became a fan of. Thus friending someone is much like subscribing to a news feed generated by a particular story submitter.

Given the amount of time it takes to find stories that have not already been submitted to Digg, one might predict that digg users would specialize: some taking the job of submitting, and others taking the job of reading and digging (voting for) the articles. On the other hand, voting on web sites is sometimes done in the context of reciprocity [2]: if you rate my story highly I will rate yours highly. This leads us to a different prediction: that those who submit will be active diggers. Also, we expect to see some evidence of notification efforts – that is, the soliciting of diggs for stories – among friendship networks. Moreover, we expect that those who are both active diggers and submitters would be more successful at promoting stories than those who focus just on submission. Even though those who focus on submission might have more time to find interesting stories, we predict that they are less likely to have a strong friendship, and thus their stories are less likely to become popular.

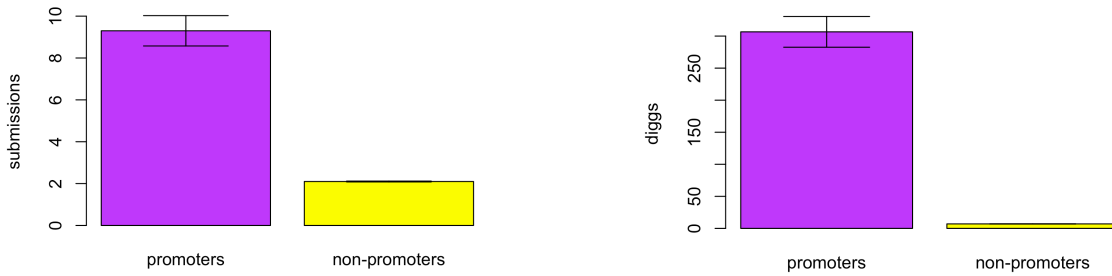
### 3. SPECIALIZATION

*Method:* We collected all story submissions on Digg from February 2008 to May 2008. We collected not just the story but the related data: the identifier of the user who made the submission, the time of submission, the set of diggs, the timing of all diggs. We picked 100,000 sample story submissions at random. About 5,000 had been removed from Digg's servers for various reasons, and so we analyzed 94,441 submissions, submitted by 42,625 distinct users. The analyses in Figures 1 through 5 are based on the submissions of these 42,625 users. Each time a story is submitted the submission counts as a digg – that is, a vote. We subtracted this initial digg, so our digg numbers are in actuality non-submittal diggs – that is, diggs by someone other than the original submitter.

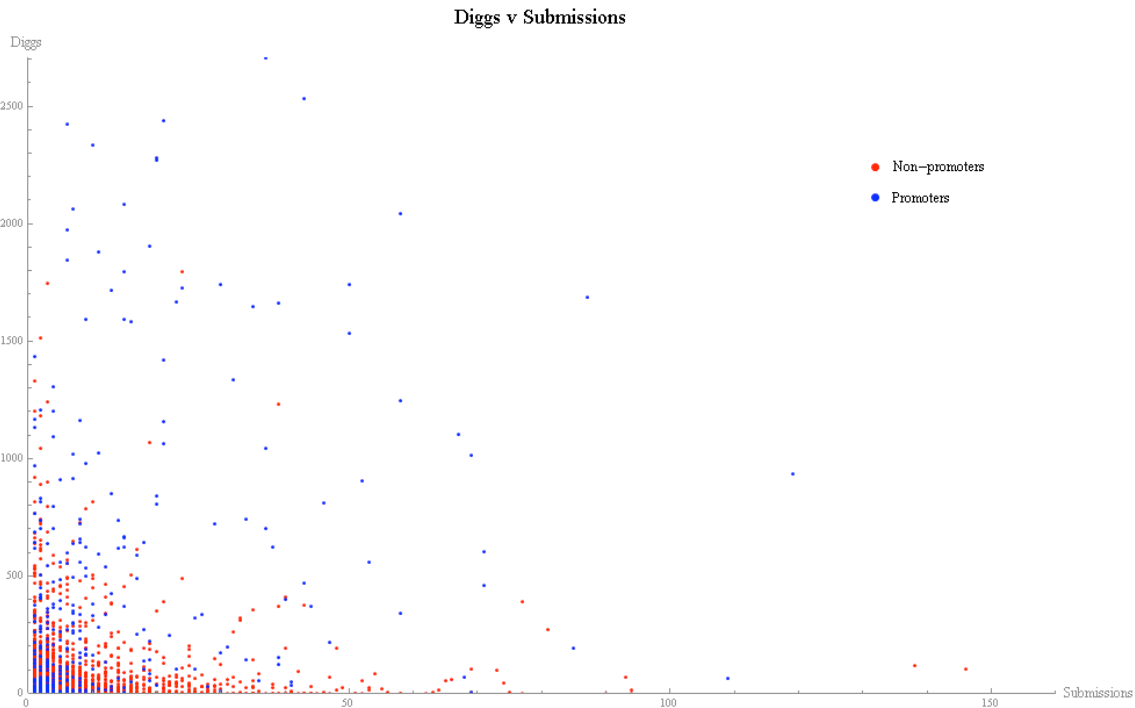
*Results:* In order to see if digging and submitting are mutually exclusive, and in order to get a sense for the nature of successful submission, we first introduced a simple classification. Most submissions never become popular – they are never promoted to popularity. Because of this, we can separate submitters into two categories: those who, in the timeframe of the study submitted at least one story that became popular (called *promoting submitters*, or *promoters* for short), and those whose stories all remained unpopular (called *non-promoting submitters*, or *non-promoters* for short). There were 42,168 non-promoting submitters, and only 457 promoting submitters. Therefore, the vast majority of submitters never see their stories promoted.

We then looked to see if the number of diggs and submissions varied across the promoting and non-promoting submitters. The means of these groups are shown in Figure 1 and Figure 2. From visual inspection of Figure 1, we can see that promoting submitters submit significantly more stories on average than non-promoting submitters. What is most surprising is the ratio of diggs to non-diggs in promoting

submitters versus non-promoting submitters: in Figure 2 promoters make far more diggs than non-promoters. From visual inspection of both sets of bar charts, we can see the differences between promoters and non-promoters are significant, and Wilcoxon rank sum tests confirm this,  $p \ll .001$ . Those whose stories get promoted are those who digg other's stories. Getting promoted is highly correlated with reading and responding to other's stories. There is a significant positive correlation between the number of promotions users have and their number of non-submittal diggs ( $r = .49, p < .001$  for promoters and  $r = .46, p < .001$  including both promoters and non-promoters).



**Figure 1.** Mean number of submissions for promoters and non-promoters. Error bars show standard error. **Figure 2.** Mean number of diggs for promoters and non-promoters.



**Figure 3.** Each point represents a single submitter over a four month time period. Blue dots indicate those whose stories have been promoted.

Figure 3 shows the data of Figures 1 and 2 together in one graph. Notice that those who digg the most, those high on the Y axis, are promoters, and that promoters dominate non-promoters along both axes – that is, the promoters, represented in blue, are the those who both submit and digg a lot, as shown by their location on the right-upper frontier of the graph. Non-promoters may have many submissions – for

example, see the red dots on the far right near the x-axis – but in general digg much less than their promoting counterparts.

The results are inconsistent with the prediction that submitters and diggers would be mutually exclusive. Instead, the results support the alternative prediction that promoters will be active diggers. However, it is not clear from the data why exactly submitters also digg so much; we analyze this in the following sections.

#### 4. NOTIFICATION MECHANISMS

*Method:* Here, we look for mechanisms through which submitters might utilize their friend networks to get stories promoted. A relatively new feature of Digg is the *shout*: users can broadcast a message to friends. These messages are only partially documented. That is, the number of people broadcast to and the time of the message are recorded, but only a subset of five recipient names are recorded. We assume that the five recorded are a random sample, and we then look to see if shouts advertising the submission of a story result in diggs from that story from recipients. We predict that some significant fraction of the notified people will digg the story. This will show a mechanism through which submitters may garner early diggs from friends to help promote a story. Such a mechanism may be important because early diggs, like the first diners of the evening at a restaurant, may quickly attract others.

We picked 22 instances of shouts at random, and checked each of the 5 identified names to see if they digg the stories.

*Results:* On average, .2 of the recipients of a shout digg the story (standard deviation = .14). The mean number of recipients of a shout was 37 (standard deviation = 18). Extrapolating, the expected number of diggs from a shout is approximately 7.

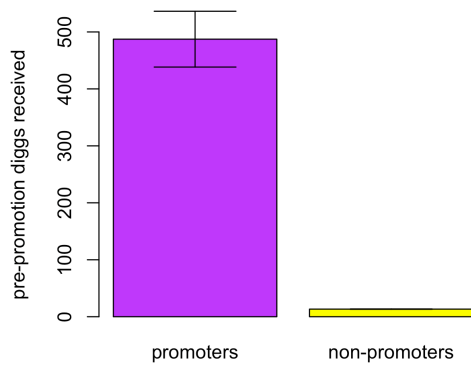
In an unexpected result, we found that many shouts did not come from the submitter, but from a third party: 42% of the shouts came from someone other than the submitter.

There are many mechanisms that might explain the role of friendship networks in Digg. Shouts are just one way that submitters can let other people know about their submission. By virtue of the Digg personalized home pages, users will see the activity of their friends. Also, diggers can email their colleagues to announce a submission. The expected number of diggs from a shout – 7 – is enough to make a difference. Most submitted stories are never digg by anyone besides the original submitter (just as, in academia, most journal articles are never cited). Therefore, a handful of initial diggs help differentiate a story, and the enrolled network of friends may help publicize the story, leading to eventual popularity.

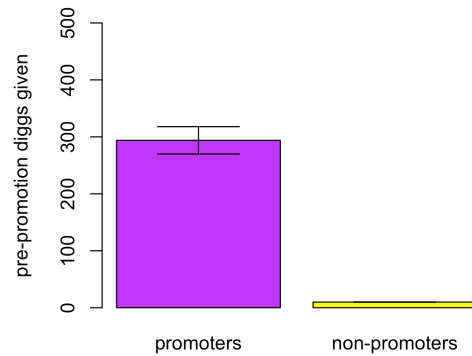
We can think of two explanations why many shout on behalf of an article they did not submit. The first is indirect reciprocation – that is, contribution without a direct expectation of reward. The second explanation is that there is a more complex form of direct reciprocation in place, and that those that shout for others will be rewarded by the submitter, who will return the favor later.

#### 5. RECIPROCITY ANALYSIS

*Method:* We define a pre-promotion digg as a digg made before a story is denoted popular by Digg. This covers all diggs for those stories that never reached popular status. The number of pre-promotion diggs a submitter receives is the sum of the number of pre-promotion diggs over all the user's submissions. Post-promotion diggs cannot of course influence popularity, because popularity has already been achieved. Thus, those who seek to make their stories popular are most interested in receiving these pre-promotion diggs. If reciprocity is at work, then we would expect that promoting submitters would both give and receive more pre-promotion diggs than non-promoting submitters. Without reciprocity, we wouldn't expect there to be significant differences between any kind of submitter, as submitting a story is an act that is independent of the act of digging someone else's story: digging a story would then be based on the story itself, not on the name of submitter or the promotion status of the submission.



**Figure 4.** Mean pre-promotion diggs *received* by promoters and non-promoters. Error bars show standard error.



**Figure 5.** Mean pre-promotion diggs *given* by promoters and non-promoters.

*Results:* From Figures 4 and 5 we can see that differences between promoters and non-promoters are significant, and Wilcoxon rank sum tests confirm this,  $p \ll .001$ . Figure 4 shows that promoters receive many more pre-promotion diggs than non-promoters. This is not unexpected – at least part of the reason that stories get promoted is because of pre-promotion diggs. What is surprising is that the promoting submitters make many more pre-promotion diggs than the non-promoting submitters. Promoting submitters are not just interested in other stories, they are almost exclusively interested in other stories before they become popular: They digg post-promotion stories at about 10% of the rate they digg pre-promotion stories.

The symmetry of Figures 4 and 5 suggest some form of reciprocity at work. We cannot at this stage in our analysis distinguish between direct and indirect reciprocity. If the reciprocity is direct, then submitters will return favors, and digg the stories of those who have dugg them in the past. If the reciprocity is indirect, then submitters actively participate in digging stories before they are promoted with the general understanding that what goes around comes around, and that their own stories will become more popular, even if there isn't direct and timely reciprocity. They simply seek to generate good Karma.

## 6. DISCUSSION AND CONCLUSIONS

Many people submit stories to Digg, but few stories become popular. The distinguishing characteristic of the successful submitters is that they are also active diggers. There are several possible explanations for this. We think that notification of story submittal happens through friend networks, and friendships are more likely to develop through reciprocal acts of digging each other's stories. But there are other possible explanations that should be tested as part of future research. Because these diggers are so active, it is possible that random effects could produce what looks like reciprocity, but is instead coincidence. Or, the appearance of reciprocity may be the result of preferences for particular story topics: if two active participants are both interested in the same things, then they may look like they are reciprocating. Finally, because some stories are more popular than other stories, it could be that active participants anticipate this popularity, and thus coincidentally digg each other's stories. All of the above alternative theories are testable, and their negation would leave reciprocation as the simplest and most plausible theory.

There is likely to be an element of indirect reciprocation – for example, we saw people soliciting support not just for their own stories, but also the stories of others. In the end, a mixture of altruistic behavior and conscious reciprocity are probably in effect in the Digg website, both aspects of embedded participation. The result is a site in which labor is donated, and in which users benefit in several ways. For the casual reader of the site, stories are filtered through a community so that popular stories are likely to represent the tastes of the users. For the active submitter, a widening set of friendship networks means

that stories posted are likely to be read, dugg, and achieve popularity, allowing the submitter an influential role.

In our previous work, we pointed out that the old marketing distinction between influentials (those who talk to people in a local community) and media stars (those who influence taste through broadcast) are blurred in online communities [11]. Here, our major contribution is a further delineation of the ways online networks create mechanisms that motivate users to contribute to a collective. Users who truly embed themselves in the site, submitting articles, and reading the articles of others, acquire influence through a network of friends. Due to the nature electronic influence, this network can grow rapidly, and provide these contributors influence on what other users will read. Thus, these influentials become local celebrities, media stars, able to broadcast ideas through their expanding network. From the standpoint of the casual user, these influentials/media stars function as editors, screening articles and publicizing those articles that will reward attention.

The nature of this community can be compared to other similar communities – for example, in academia, those who write highly cited articles are likely to also read, edit, and cite articles written by others. In all such communities, the presence of reciprocity is problematic. On the one hand, it is possible that work gets promoted more because of the social network of the submitter than because of the intrinsic merit of the work. On the other hand, these networks of reciprocity are highly motivating, and encourage participants to maintain an awareness of the community that surround them. In the end, communities of practice are always a combination of motivating mechanisms, social networks, and artifacts; the interpretation of artifacts will always be affected by the community the interpreters are embedded within.

## REFERENCES

- [1] Tovey, M. (ed.) *Collective Intelligence: Creating a Prosperous World at Peace*, Earth Intelligence Network, 2008.
- [2] Dellarocas, C. N., Fan, M. N. and Wood, C. A., Self-Interest, Reciprocity, and Participation in Online Reputation Systems (February 2004). MIT Sloan Working Papers No. 4500-04.
- [3] Nowak M. A, K Sigmund. Evolution of indirect reciprocity by image scoring. *Nature* 393: 1998, 573-577.
- [4] Montgomery, J. D. "Toward a Role-Theoretic Conception of Embeddedness," *American Journal of Sociology*, 104 (July 1998): 92-125.
- [5] Granovetter, M. "Economic Action and Social Structure: The Problem of Embeddedness"; *American Journal of Sociology*, Vol. 91, No. 3., November 1985, 481-510
- [6] Mohtashemi M., Mui L. Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altruism. *Journal of Theoretical Biology* 223(4), 2003, 523-31.
- [7] Fehr, Ernst; and Simon Gächter "Fairness and Retaliation: The Economics of Reciprocity". *Journal of Economic Perspectives* 14 (3): 2000, 159–181.
- [8] Hollan, J., Edwin Hutchins, and Kirsch, D. Distributed Cognition: Toward A New Foundation for Human-Computer Interaction Research, *ACM Transactions on Computer-Human Interaction*, Volume 7, Issue 2, 2000, 174 -196.
- [9] Sahlins, M. *Stone Age Economics*. Chicago: Aldine-Atherton. 1972.
- [10] Cialdini, R. B. *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon. 2001.
- [11] Sakamoto, Y., Sadlon, E. and Nickerson, J. V. Bellwethers and the Emergence of Trends in Online Communities, *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, 2008.