The Ecology of Digg:
Niches and Reciprocity in a Social Network Landscape

Elliot Sadlon
Stevens Institute of Technology

Yasuaki Sakamoto
Stevens Institute of Technology

Jing Ma
Stevens Institute of Technology

Samuel Barrett
University of Texas at Austin

Jeffrey V. Nickerson
Stevens Institute of Technology

1. Introduction

Web sites such as Digg and Wikipedia, which rely on user participation, are quite successful in continuously providing popular services. These participation-driven sites are examples of collective intelligence (e.g. [1]). In the case of Digg, for example, users determine which stories are important by acting as editors or filtering devices for other users.

How do participation-driven web sites self-organize? Examinations of participation-driven sites reach different conclusions. In the extreme, whereas some think that the participant-driven web sites are basically the result of altruistic impulses [1], others think that they are the result of selfish impulses, in which participants are mainly interested in building their own popularity through reciprocity [2]. As with any social phenomenon, it is probable that the web sites have mixed populations of individuals with different preferences and motivations.

In the hopes of answering in part the question of how the site self-organizes, we extend theories from natural ecology, human ecology, and evolutionary biology to human social behavior using Web-based technologies. In such theories, actors with different capabilities compete and cooperate, striving for survival in a landscape within which resources may be limited. Moreover, situations change, actors adapt, and structure evolves.

Here, we apply this ecological metaphor to a web environment, in which participants do much of the labor for a web site on a volunteer basis. Such sites include SourceForge, Wikipedia, and Digg. In this paper we will focus on Digg, the least studied of the three (although see [20], [21], [23]). In a nutshell, participants in Digg submit links to news stories or other web-based content, and other participants may rate the stories up or down (*digg* or *bury*, in the terminology of the site), and may also comment on the stories. Based on the history of Diggs and buries, about 1% of the stories submitted are declared by the web site owners to be popular, usually within 24 hours of being posted, and that declaration leads to a prominence of display on the site that in turn leads to many more participants reading the story. Thus, the site collectively plays the role of a newspaper editor, deciding which stories will run on the first page.

We think that Digg, and sites like it, can be characterized as an ecology. We will look at two aspects of the site that are distinct – the specialization of users around story topics, and the reciprocity displayed in votes related to submitted stories. These aspects illustrate different aspects of competitive and cooperative behavior on the site, and together give us more nuance look at the site's organization than either can alone.

These two kinds of phenomenon are of interest to many different disciplines, including information systems, management, anthropology, and psychology. Specialization in particular has been studied as part of natural ecology, and the extensions of natural ecology into the study of human behavior (e. g. [1-7]). Reciprocity, defined here as an action taken in the expectation of a return action in kind, has been studied in a field related to ecology, such as evolutionary biology [8-10], social/evolutionary psychology [11-14], and behavioral economics [15-17]. These fields provide a way of thinking about social networks, a metaphor in which we can speak of specialization, niches, survival, reciprocity, and altruism. The relational and dynamic natures of ecologies are analogous to those of participant-driven websites, and thus we can utilize these theories to make predictions that can be tested empirically.

In particular, the ecological account predicts that users compete for their metaphoric survival with regard to story topic. Users gain status in Digg (and, following the metaphor, prolong survival) by establishing a reputation through the submission of stories that later become popular. Competition among users for topic space is advantageous for the Digg community because more space will be covered thus providing viewers of Digg access to information from a wider variety of sources.

The ecological account also predicts that users who are successful at surviving will be those who reciprocate (e.g., [11]). In social psychology, reciprocal actions reward positive actions with other positive actions and punish negative actions with negative actions, and are important for maintaining social norms, or a set of behavioral expectations within a social group, which in turn are critical for the survival of the group [13]. Positive reciprocity (e.g., helping a user by digging that user's story) increases the chance of future return (e.g., that user diggs back), and such behavior can increase the rate of overall contribution from the users as a whole [15].

In this paper, we first demonstrate that a critical set of users, the submitters, specialize with respect to story topics. That is, they compete for resources and form niches. Then, we show that users do not specialize with respect to submitting and voting on stories; instead, they form reciprocal relationships, with a small set of users both submitting stories and also voting on other users' stories far before the general public reads them, to help their stories gain popularity. This reciprocity, we argue, is probably not related to story topics, but instead is a result of an awareness of the social network. In other words, decisions to digg a story are socially mediated.

2. TOPIC LANDSCAPE

Method

For the first analysis we used data on the top 100 submitters and their most recent 100 submissions taken on July 2, 2008. We looked at the 51 topics that Digg provides for users to submit stories into. For each topic, we calculated the frequency of the words in the topic by totaling the word frequencies of each story submitted in that topic. Then, we calculated a distance matrix containing the distances between each two Digg topics based on the city-block distances between their respective word frequencies. Finally, we used multi-dimensional scaling (cf. [22]) to convert the pair-wise topic distance matrix into a two-dimensional topic space that preserves as closely as possible the word-frequency distances between topics.

Results

Figure 1 is the resulting multidimensional scaling diagram showing the Digg topic landscape. Such diagrams can be interpreted in different ways [22]. Our interpretation is that the vertical axis runs from politics on the top (general and people oriented) to hardware (specific and machine oriented) on the bottom. The horizontal axis runs from sports on the far left to general and tech news on the far right. The two most popular topics on Digg, *tech news* and *odd stuff,* are both close together on the mid-right of the diagram in a large cluster of similar general news topics. Given the Digg topic landscape, our question is how story submitters position themselves in this landscape.

Figure 1. Multidimensional scaling of Digg topics, based on a similarity matrix built from textual analysis of the stories submitted by the top diggers.

3. SUBMISSIONS BY DIGG TOPIC

Method

In our second analysis, we looked at submissions by topic over a several month period. We took the top 100 submitters and analyzed their topic submissions. Our ecological account predicts that there will be a specialization by topic. For our analyses, we concentrated on 34 of the most popular topics, and categorized the submitters by the number of these topics that they dominated – that is, the topics that they submitted more stories for than the other submitters.

Results

Figure 2 shows the dominating submitters and the topics they dominate. For example, in the first row is *mrbabyman*, who dominates seven topics (blue indicates topic dominance). The second and third rows also show submitters, *bonlebon* and *mklopez*, who are able to dominate across a wide range of topics.

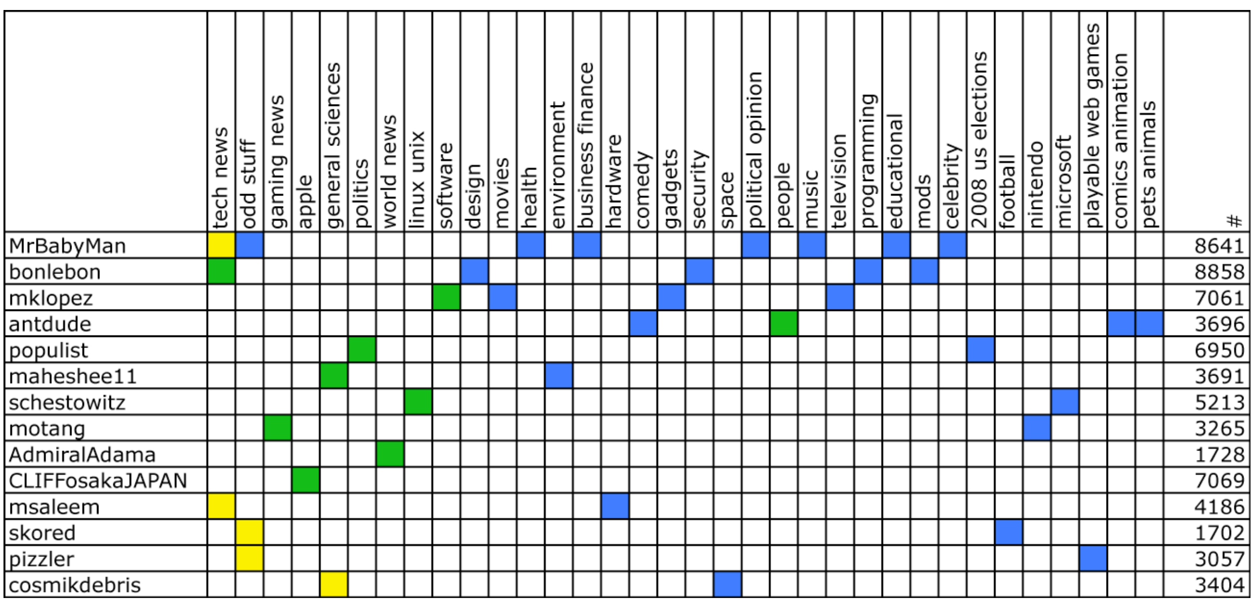

Figure 2. Across the top, the 34 most submitted-to topics. Along the side, the 14 submitters who dominate at least one topic category. Blue indicates category dominance. Yellow indicates the most submitted-to category for the individual. Green indicates the intersection – categories that were the most submitted to by an individual, and also dominate the category. The numbers to the right are the total submissions for each subject in these 34 most popular categories.

The rest of the submitters shown here appear to specialize. For example, *schestowitz* dominates the *Linux* category, submitting 56% of the total submissions in a group of 74 submitters. Of his own submissions, 86% are to this category. He also dominates the *Microsoft* category – we can infer that he is specializing in operating system-related news. For other submitters, we can even see specialization reflected in the names submitters chose for their handles: *populist* dominates *politics*, *skored* dominates *football*, and *cosmikdebris* dominates *space*. Both a chi-squared test and a Cochran Q tests showed the distributions of dominance were not uniform, $p < .001$, suggesting that this specialization is unlikely to occur by chance.

Figures 1 and 2 are related. If we look at those who dominate several topics, we notice that often these topics are also close together in Figure 1. For example, *populist* submits to both *politics* and the *2008 US election*. *Schestowitz* submits to *Linux* and *Microsoft*. *Motang* submits to *gaming news* and *Nintendo*. Thus, it appears that the distances shown in Figure 1 are also sensed by the submitters, and, for the most part, submitters tend to focus their effort in topics that are close together.

4. RECIPROCITY

Participants have different preferences, and self-organize in relationship to topics. According to the ecological account, the specialization with respect to topic is important for creating more popular stories. The ecological account also predicts that reciprocity plays an important role in producing more popular stories in Digg [23]. To reciprocate, users need to both submit and digg, instead of specializing on one action. In addition, reciprocity leads to some users digging each other's stories and clustering together.

Method: We collected all story submissions on Digg from February 2008 to May 2008. We analyzed 94,441 random submissions, submitted by 42,625 distinct users.

Results: Most submissions never become popular, because they are never promoted to popularity. As a result, 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: Most submitters never see their stories promoted.

We then examined the number of diggs and submissions varied across the promoting and non-promoting submitters. The means of these groups are shown in Figure 3 and Figure 4. As expected, Figure 3 shows that promoting submitters submit significantly more stories on average than non-promoting submitters. More interesting, Figure 4 shows that promoters make far more diggs than non-promoters. 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). There results suggest, as predicted by the ecological account and the idea of reciprocity, that successful users contribute to the community by both finding stories (submitting) and acting as editors (digging).

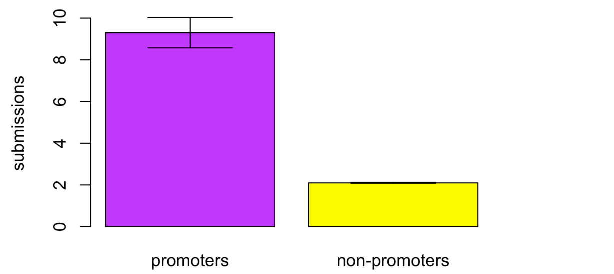

Figure 3. Mean number of submissions for promoters and non-promoters. Error bars show standard error.

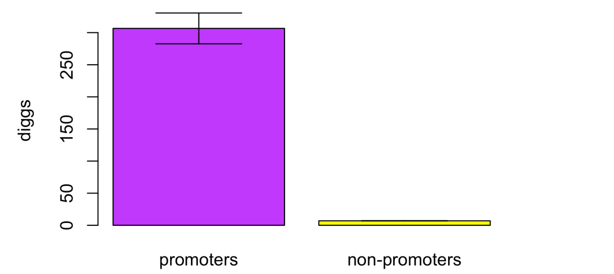

Figure 4. Mean number of diggs for promoters and non-promoters.

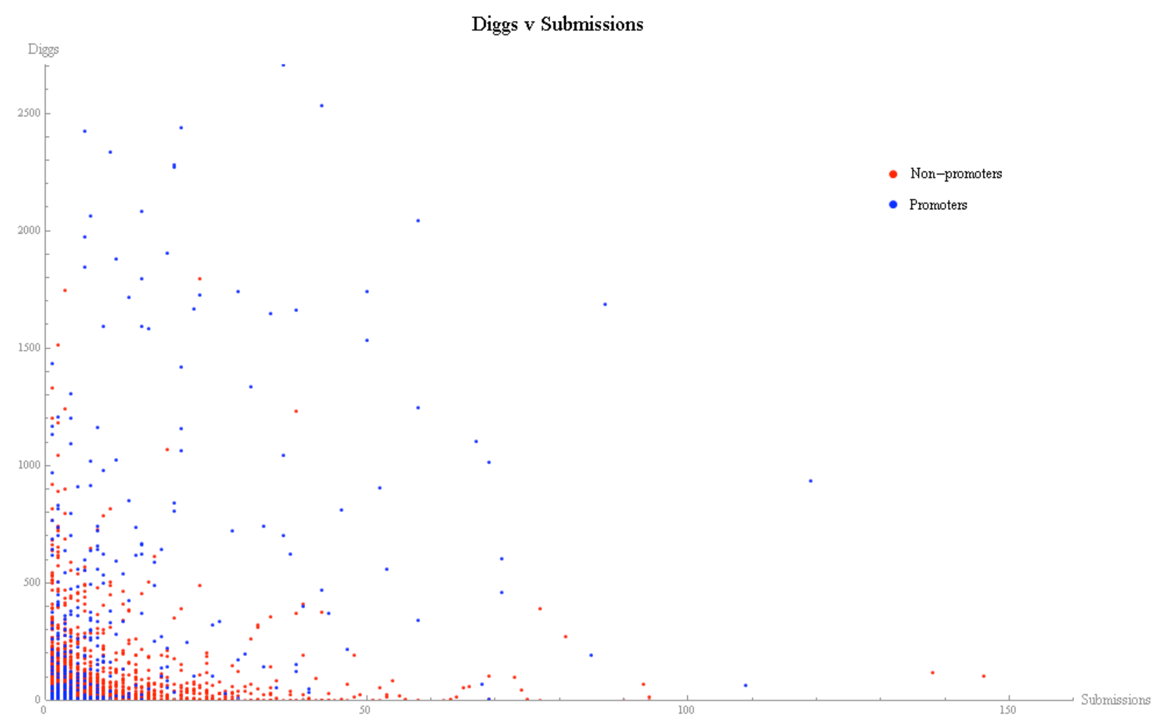

Figure 5. Each point represents a single submitter over a four month time period. Blue dots indicate those whose stories have been promoted.

Figure 5 shows the data of Figures 3 and 4 in a different way. Those who digg the most, those high on the Y axis, are promoters, and promoters dominate non-promoters along both axes: 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 but digg much less than their promoting counterparts.

The results support the prediction that promoters will be active diggers. In our analysis in [23] we found that submitters and diggers have ways to signal each other, and at least some appear to be reciprocating, consistent with findings in other online environments (e.g. [2], with respect to eBay). We have recently ruled out several alternative hypotheses through simulation modeling: the appearance of reciprocity is not due to randomness, nor to preference for topic, nor to the popularity of the stories.

We further examined more direct evidence for reciprocity. If users are conscious of reciprocity, then they probably keep the number of diggs they give roughly equal to those they receive. To look for such patterns, we created a matrix shown in Table 1: users listed in the rows received diggs from users listed in the columns. We then created a scatter plot of user pairs as shown in Figure 6. For example, Emberjohn, has roughly symmetric diggs with 1KrazyKorean. As displayed in Figure 6, there is a group of users that have roughly symmetric diggs with each other – those that are located close to the diagonal line. In addition, in Figure 6, user pairs far from the diagonal line, at the lower right corner, may represent instances of a leader-follower relationship. For instance, jjw269 gave a large number of diggs to reflex768, mpmind176, and MookiBlaylock, but received only a few diggs from those users in return. These three receiving users may be trendsetters, and Jjw269 may be following their lead. Figure 6 provides a way to easily identify both reciprocity and leader-follower relationships.

Table 1. The number of diggs between a subset of users. Rows received diggs from columns.

	mpind176	oboy	Bukowsky	MookiBlaylock	reflex768	1KrazyKorean	kickoff22	emberjohn	cmccool	bcuban	skateboard1	Inthenameofmine	ohnoerino	jjw269	theeandrew	iraq	MississippiLife	jstohler	Hawker400	casspa	dan360man	orlando69	queenmoweeny
mpind176		26	13	66	26	21	13	25	0	68	0	44	0	62	80	62	0	3	5	28	0	0	4
oboy	53		28	51	67	17	6	28	4	9	46	39	4	6	32	44	18	3	5	27	2	10	13
Bukowsky	69	46		63	48	33	7	27	2	15	9	30	2	3	31	43	0	4	2	21	0	15	8
MookiBlaylock	74	36	8		25	16	7	7	2	19	2	51	5	71	80	4	0	1	3	21	0	37	0
reflex768	36	50	4	12		12	2	20	0	1	47	37	0	65	0	5	14	0	0	16	0	2	1
1KrazyKorean	22	30	48	13	21		1	37	1	6	13	13	0	9	20	4	7	1	0	27	1	7	3
kickoff22	39	3	9	18	24	3		14	1	32	0	2	14	11	3	3	0	8	4	27	0	20	1
emberjohn	2	22	17	0	14	36	4		0	0	11	1	0	21	8	19	8	0	1	2	0	0	11
cmccool	4	30	8	14	6	0	1	33		1	10	5	0	0	1	8	0	3	0	22	0	1	0
bcuban	28	7	0	21	5	5	12	10	1		1	10	3	5	13	16	0	4	2	11	1	9	0
skateboard1	4	30	2	0	37	2	0	16	0	0		0	0	28	0	10	0	1	0	3	0	0	8
Inthenameofmine	6	6	4	3	4	2	4	14	0	5	0		0	19	7	16	1	2	1	5	0	6	0
ohnoerino	5	6	2	15	2	0	12	1	3	9	0	5		0	1	0	0	0	3	17	16	1	0
jjw269	4	1	0	4	11	1	1	10	0	0	7	12	0		2	5	0	4	0	0	0	2	0
theeandrew	10	6	2	6	1	6	0	4	0	3	0	5	0	2		8	0	3	2	4	0	4	0
iraq	1	11	3	0	0	0	0	14	0	0	9	10	0	7	1		0	1	0	0	0	0	1
MississippiLife	0	13	3	0	12	1	4	9	0	0	0	5	0	0	0	1		1	0	3	0	0	1
jstohler	0	0	0	0	15	1	0	3	0	0	0	0	0	6	0	0	0		3	1	0	0	1
Hawker400	2	3	2	3	0	1	1	1	0	1	0	0	0	0	0	1	0	19		1	0	0	0
casspa	2	3	0	2	4	4	2	5	1	2	0	0	2	0	0	1	0	0	0		0	0	0
dan360man	2	0	0	1	1	0	0	3	1	0	0	9	1	0	0	0	0	1	0	4		0	0
orlando69	1	2	0	2	0	0	0	0	0	1	0	4	0	0	6	6	0	1	1	2	0		0
queenmoweeny	0	5	0	1	3	0	0	5	0	0	5	0	0	0	0	3	1	0	0	0	0	0	

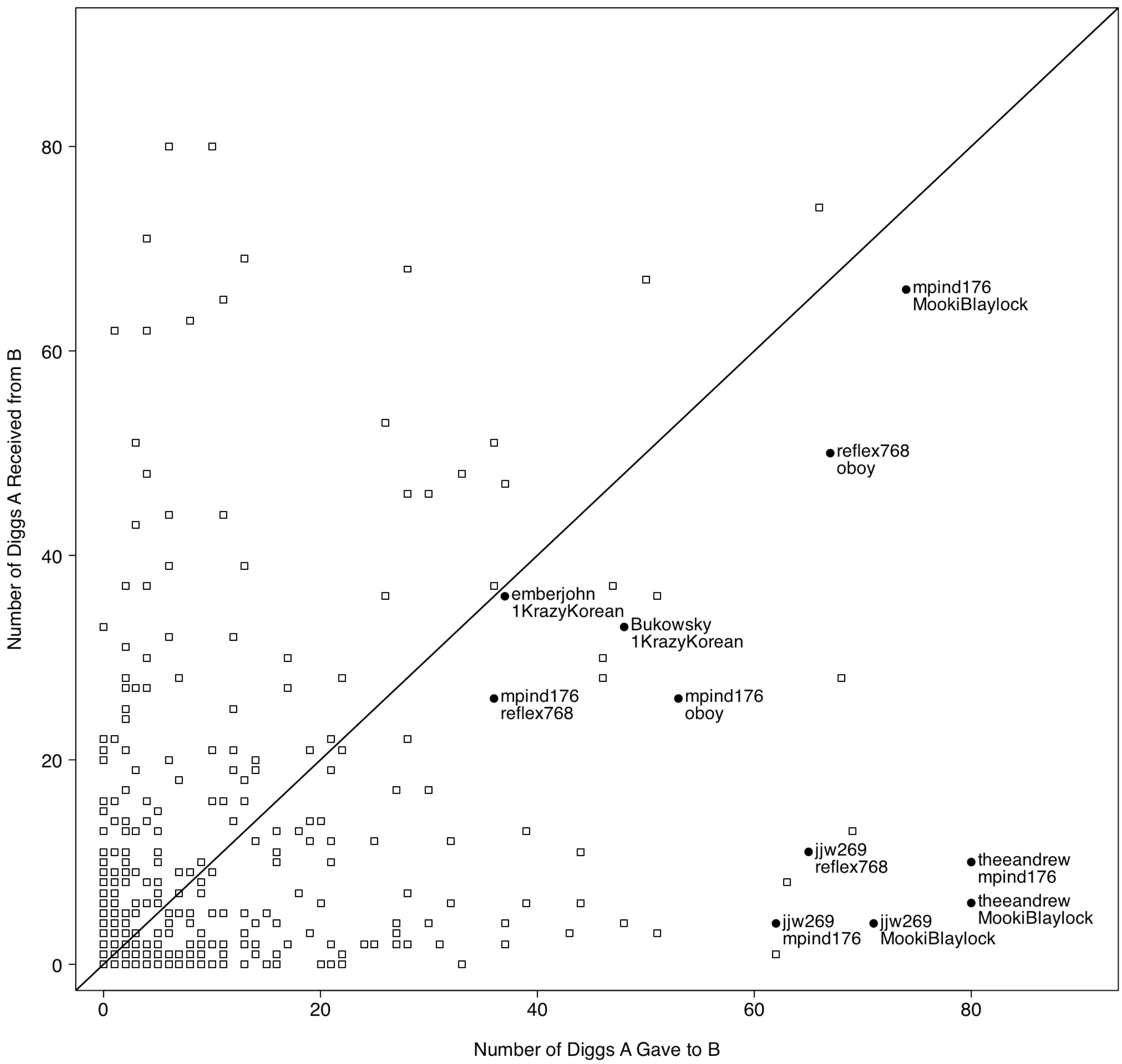

Figure 6. A scatter plot visualization of the data in Table 1. For example, Emberjohn gave 37 Diggs to and received 36 Diggs from 1KrazyKorean. Several of the interesting points are labeled. The closer to the midline, the more symmetrical the relationship. The closer to the right top corner, the more activity between the pairs of users. Points on the bottom right may indicate leader-follower relationships.

Using the matrix in Table 1, we propose the following measure of reciprocity similarity:

(diggs made * diggs received) / (diggs made + diggs received).

This can be thought of as an area divided by a perimeter: given a particular number of diggs, the measure is at a maximum when the relationship is perfectly reciprocal. A pair (10,10) will in this metric be more reciprocal than a pair (28, 2), in agreement with our intuition. Likewise, a pair (20, 20) will be seen as more reciprocal than a pair (10,10). In Table 2, we show the results of this calculation for a subset of the users shown in Table 1.

Table 2. Reciprocity measure between users.

	oboy	Bukowsky	MookiBlaylock	reflex768	X1KrazyKorean	kickoff22	emberjohn	cmccool	bcuban
mpind176	17.4	10.9	34.9	15.1	10.7	9.8	1.9	0.0	19.8
oboy		17.4	21.1	28.6	10.9	2.0	12.3	3.5	3.9
Bukowsky			7.1	3.7	19.6	3.9	10.4	1.6	0.0
MookiBlaylock				8.1	7.2	5.0	0.0	1.8	10.0
reflex768					7.6	1.8	8.2	0.0	0.8
1KrazyKorean						0.8	18.2	0.0	2.7
kickoff22							3.1	0.5	8.7
emberjohn								0.0	0.0
cmccool									0.5

After converting this similarity measure into a distance measure, we can then cluster this set of users, as shown in figure 7. In the center of the diagram is a set of users, 1KrazyKorean and Bukowsky. We looked at cycles between users in a bipartite graph of users and stories – that is, instances in which user *A* diggs a story of *B*'s, and then *B* diggs a story of *A*'s. 1KrazyKorean and Bukowsky occurred in 21 cycles together: mpind176 and MookiBlaylock are also clustered together in Figure 7. These two users appeared together in 21 direct cycles. Similarly, Oboy and reflex768, who cluster in Figure 7, appeared in 12 cycles together. Our clustering analysis based on the reciprocity measure correctly identifies those individuals who participate in cycles.

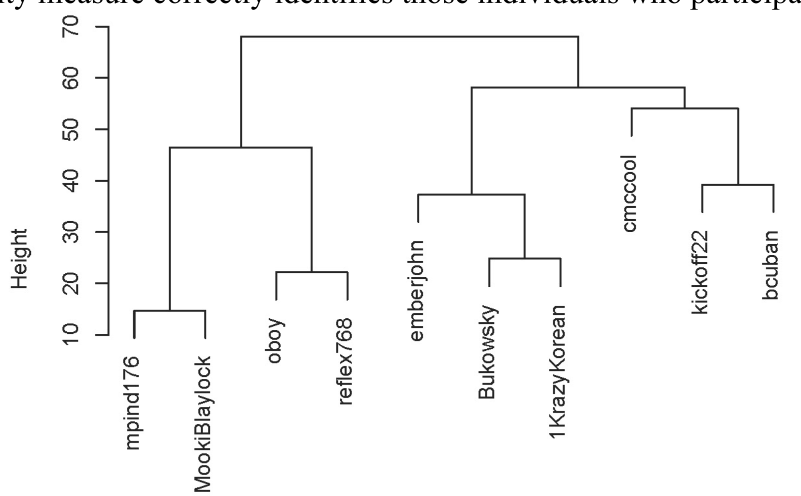

Figure 7. Hierarchical cluster based on the reciprocity measure.

Interestingly, the matrix scatter plot in Figure 6 is consistent with the clustering of reciprocating users. The three pairs, 1KrazyKorean-Bukowsky, mpind176-MookiBlaylock, and Oboy-reflex768, identified in the cluster are all located close to the diagonal line, indicating the pairs have a roughly symmetric relationships with regard to the number of diggs given and received.

We also looked at larger sized cycles in the bipartite graph of users and stories. There was only one cycle involving any three users that occurred more than once: 1KrazyKorean, Bukowsky and emberjohn participated in 9 such cycles. These users are placed adjacent to one another in the hierarchical cluster in Figure 7.

Figures 8 and 9 visualize the identified clusters in a different way. Figure 8 chooses one user, mpind176, and shows his relations to the three other users in the cluster on the left of figure 7. Two out of the three pairs so closely matching diggs given and received. Figure 9 shows the two cluster neighbors of 1KrazyKorean; the relationship to Emberjohn is almost perfectly symmetrical, and the relationship to Bukowsky is also quite strong: each is digging many stories of the other. At the very least, the graphs suggest that these users are conscious of what the others are doing.

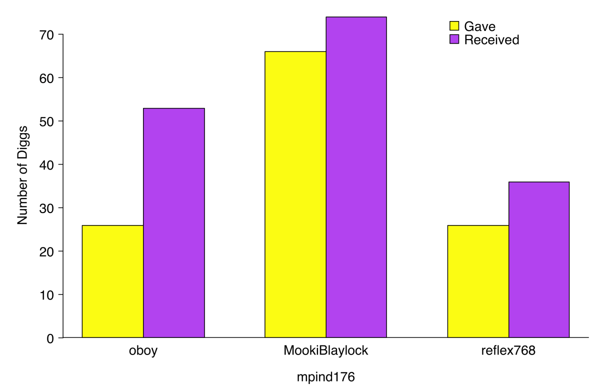

Figure 8. An example of the reciprocal relationships for the subset of users identified on the left side of cluster in Figure 7. These are the relations of mpind176.

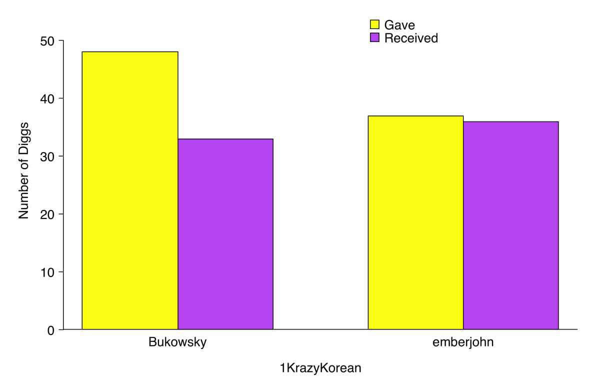

Figure 9. Another example of the reciprocity relationships for the subset of users identified in the center of Figure 7. These are the relations of 1KrazyKorean.

It is unclear whether or not cycles occur because participants digg each other often just because they respect each other's taste, or whether participants seek to close open cycles (obligations), and therefore digg each other a lot. What is clear is that there are several pairs of users who seem to digg each other's work symmetrically.

There is another group of users who seem to give more than they receive. Such users are evident at the lower right corner in Figure 6. For instance, jjw269 and theeandrew each gave a large number of diggs to mpmind176 and MookiBlaylock, but received only a few diggs from those users in return. These users may be following the behavior of influential users. Future research might explore whether or not those that give much more than they receive are users who are relatively junior in experience and reputation, seeking to climb the reputation ladder.

5. CONCLUSIONS

There is a landscape of news stories, and submitters of these stories tend to specialize in particular areas. In ecological terms, they find a niche. The topic niches, however, do not explain how the site works. 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. It appears that some set of diggers practice a form of reciprocity: diggs are awarded with return diggs. Participants create a system useful to many of us while, it seems, they engage in the same kinds of competitive and cooperative practices that define much of our cultural coexistence.

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*, MIT Sloan Working Papers No. 4500-04, February 2004.

[3] Hawley, A. H. *Human Ecology: A Theoretical Essay*, Chicago: Chicago University Press, 1986.

[4] Baum, J. A. C. and Oliver, C. Toward an institutional ecology of organizational founding, *Academy of Management Journal,* 39, 5, 1996, 1378-1427.

[5] Baum, J. A. C. and Oliver, C. Institutional Embeddedness and the Dynamics of Organizational Population, *American Sociological Review*, 57, 4, 1992, 540-555.

[6] Goldstone, R. L., Jones, A., and Roberts, M. E. Group path formation, *IEEE Transactions on System, Man, and Cybernetics* , 36, 2006, 611-620.

[7] Gureckis, T. M., and Goldstone, R. L. Thinking in groups in Distributed cognition, *Special Issue of Pragmatics & Cognition*, in S. Harnad and I. Dror (Eds.), Amsterdam, The Netherlands: John Benjamins, 14, 2006, 293-311.

[8] Lieberman, E., Hauert, C., and Nowak, M. A. Evolutionary dynamics on graphs. *Nature,* 433, 7023, 2005, 312-316.

[9] Nowak, M. A., and Sigmund, K. Evolution of indirect reciprocity by image scoring. *Nature,* 393, 6685, 1998, 573-577.

[10] Ohtsuki, H., Hauert, C., Lieberman, E., and Nowak, M. A. A simple rule for the evolution of cooperation on graphs. *Nature,* 441, 7092, 2006, 502-505.

[11]] Axelrod, R. *Evolution of Cooperation*, New York: Basic Books, 1984.

[12] Robert, K., McCabe, K., Smith, V., and Wilson, B., Incremental Commitment and Reciprocity in a Real Time Public Goods Game, *Personality and Social Psychology Bulletin*, 27, 12, 2001, 1662-1673.

[13] Aronson, W. A. *Social Psychology* (6th Ed). New Jersey: Pearson Education, Inc, 2007.

[14] Cialdini, R. B. *Influence*, New York, NY: Morrow, 1984.

[15] Fehr, E. and Gächter, S. Fairness and Retaliation: The Economics of Reciprocity*, Journal of Economic Perspectives*, 14 , 3, 2002, 159–181.

[16] Fehr, E., Fischbacher, U., and Gächter, S. Strong reciprocity, human cooperation, and the enforcement of social norms, Human Nature, 13, 2002, 1-25.

[17] Hoffman, E., McCabe, K., and Smith, V., Behavioral Foundations of Reciprocity: Experimental Economics and Evolutionary Psychology, *Economic Inquiry*, 36, 3, 1998, 335-52.

[18] Hannan, M. T., and Freeman, J. The Population Ecology of Organizations, *American*

Journal of Sociology, 1977
[19] Nickerson, J.V., and zur Muehlen, M. The Ecology of Standards Processes: Insights from Internet Standard Making," MISQ, (30: Special Issue, August), 2006, 467-
[20] Lerman, K., User Participation in Social Media: Digg Study, *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, 2007, 255-258.
[21] Sakamoto, Y., Sadlon, E. and Nickerson, Jeffrey V. Bellwethers and the Emergence of Trends in Online Communities, Proceedings of the 30th Annual Conference of the Cognitive Science Society, 2008.
[22] Kruskal, J., and Wish, M. *Multidimensional Scaling*, Sage Publishing, Newbury Park, CA, USA, 1978.
[23] Sadlon, E., Sakamoto, Y, Nickerson, J. V., The Karma of Digg: Reciprocity in Online Networks, Workshop on Information Technologies and Systems (WITS), December 2008 (in press)