

---

# Dynamics and diversity of online community activities

T. HOGG<sup>1</sup> and G. SZABO<sup>1</sup>

<sup>1</sup> *HP Labs - Palo Alto, CA*

PACS 89.75.Hc – Networks and genealogical trees

PACS 89.20.Hh – World Wide Web, Internet

PACS 87.23.Ge – Dynamics of social systems

**Abstract.** - Web sites where users create and rate content as well as form networks with other users display long-tailed distributions of user activity. Using data from one such community site, Essembly, we propose and evaluate mechanisms for these distributions that rely only on information actually available to users. For Essembly, we describe the scale-free degree distribution of the social network as a result of users' shared interests, manifested by their content rating activity. We find the long tails in network properties arise from user activity rates that are broadly distributed, as well as the extensive variability in the time users devote to the site.

---

**Introduction.** – Participatory web sites facilitate their users creating, rating, and sharing content. Examples include Digg[.com] for news stories, Flickr[.com] for photos, and Wikipedia[.org] for encyclopedia articles. To aid users in finding content, many such sites employ collaborative filtering [1] to allow users to specify links to other users whose content or ratings are particularly relevant. These links can involve either people who already know each other, or people who discover their common interests through participating in the web site. In addition to helping identify relevant content, the resulting networks enable users to find others with similar interests and establish trust in recommendations [2].

The availability of activity records from these sites has led to numerous studies of user behavior and the networks they create. Observed commonalities in these systems suggest general generative processes leading to these observations. Examples include preferential attachment in forming networks and multiplicative processes leading to wide variation in user activity. Moreover, observed behavior can arise from a variety of mechanisms [3, 4].

Identifying information readily available to users on a participatory web site can suggest plausible causal mechanisms. The simplest such approach considers average behavior of users on a site [5]. Such models can indicate how system behavior relates to the average decisions of many users. By design, such models do not address a prominent aspect of observed online networks: the long tails in their distributions of links and activity. Models including this diversity could be useful to improve effectiveness of

the web sites by allowing to focus on significantly active users or especially interesting content, and enhancing user experience by leveraging the long tail in niche demand [6].

A key question with respect to the observed diversity is whether users and the networks are reasonably viewed as behaviors arising from a statistically homogeneous population, and hence well-characterized by a mean and variance. Or is diversity of intrinsic characteristics among participants the dominant cause of the observed wide variation in behaviors? Moreover, to the extent user diversity is important, what is a minimal characterization of this user variation sufficient to produce the observed long-tail distributions?

This paper considers these questions in the context of a politically-oriented web community, Essembly<sup>1</sup>. Unlike most such sites, Essembly provides multiple networks with differing nominal semantics, which is useful for distinguishing among some models. We consider plausible mechanisms users could be following to produce the observed long-tail behaviors both in their online activities and network characteristics. In the remainder of this paper, we first describe Essembly and our data set, then separately examine the network formation of users and then their highly variable behaviors, respectively. We suggest models to describe the observed characteristics of the network and the users.

**Essembly.** – Essembly is an online service helping users engage in political discussion through creating and voting on *resolves* reflecting controversial issues. User

---

<sup>1</sup>Essembly LLC at <http://www.essembly.com>

activities consist of *creating resolves* (users can post resolves publicly), *voting* (expressing their opinions on resolves on a 4-point scale ranging from *strong disagreement* to *strong agreement*), and *forming links* (linking to other users through three kinds of links, see next).

Essembly provides three distinct networks for users: a social network, an ideological preference network, and an anti-preference network; called *friends* (those who know each other in person), *allies* (those who share similar ideologies), and *nemeses* (those who have opposing ideologies), respectively. The type of connection is indicated by the users when they create a link on the web interface. The resulting networks have structure similar to that seen in other social network web sites [7].

The distinct social and ideological networks enable users to distinguish between people they know personally and people encountered on the site with whom they tend to agree or disagree. Network links are formed by invitation only and each link must be approved by the invitee. Thus all three networks in Essembly are explicitly created by users. Essembly provides on each user’s personal page a ranked list of ideologically most similar or dissimilar users based on voting history, thus users can identify potential allies or nemeses by comparing profiles. With regards to voting activity, the Essembly user interface presents several options for users to discover new resolves, for instance based on votes by network neighbors, recency, overall popularity, and degree of controversy.

The data set of 15424 users consists of anonymized voting records for Essembly between its inception in August 2005 and December 2006, and the users and the links they have in the three networks at the end of this period. Essembly presents 10 resolves during the user registration process to establish an initial ideological profile used to facilitate users finding others with similar or different political views. To focus on user-created content, we consider the remaining 24953 resolves, with a total of 1.3 million votes.

**Links.** – Users’ decisions of who to link to and how they attend to the behavior of their neighbors can significantly affect the performance of participatory web sites. A common property of such networks is the wide range in numbers of links made by users, characterized by the degree distribution of the network. The structure of the three networks in Essembly is typical of those seen in online social networking sites: the degree distributions are close to a truncated power law [8], with the number of users in the network with degree  $d$  proportional to  $d^{-\tau} e^{-d/\kappa}$ . Fig. 1 shows the distribution of degrees in the networks.

These long-tailed degree distributions are often viewed as due to a preferential attachment process, which, combined with a limitation on the number of links a user has, gives truncated power-law degree distributions [9,10]. However, users in Essembly have no direct access to the number of links of other users, thus we need to identify

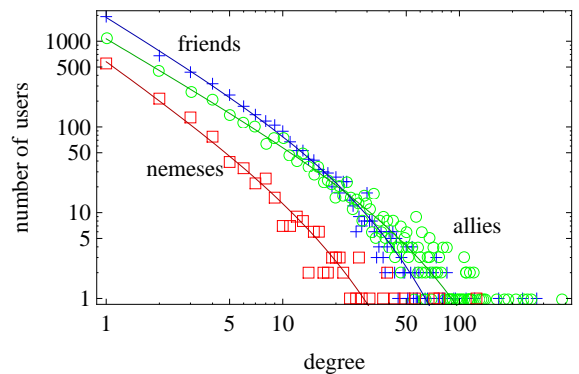


Fig. 1: The number of users in the Essembly social network who have a given number of links of the indicated type (plus symbols are for the friends, circles for the allies, and squares for the nemeses networks, respectively). The parameters of the best fits of truncated power laws on the three sets of data are given in the text.

a mechanism users could use, based on information available to them. It is plausible to assume that Essembly users find each other through shared interest in societal matters, and establish links not considering how many connections they already have (as would be the case in structure-based models), but instead on similarities between their voting activities. In particular, we show in the following that hidden-variable models may give rise to the observed degree distribution of the social and ideological networks.

The number of links a user forms is a combination of his activity rate and how long the user remains at the site. The wide variation in activity times and rates among users (see Fig. 5 and 6 later) give rise to a wide distribution of the number of links. While the most common mechanisms designed to reproduce the observed power-law degree distributions in most natural networks use growth rules and the degree of vertices in link formation [11], in the following we propose a mechanism that only takes into account the extent to which two users share interests to describe link formation between two users.

We take the friends network to primarily reflect a pre-existing social network, one that does not arise as a result of users’ activity on the website. We ignore the possibility that users who do not know each other in real life can also form friendship links, since this is explicitly discouraged by the site rules presented at signup. For the ally/nemesis networks, however, we take the linking choices to depend on the voting activity of the users forming connections. This assumption is motivated by the fact that Essembly prominently highlights ideologically similar and dissimilar users on one’s personal page, which is a list that is inferred from votes on the same resolves by two users. This makes it particularly easy for users to make connections with people who share many common votes.

People thus most likely form ideological links based on similarities in their voting patterns, and only users active

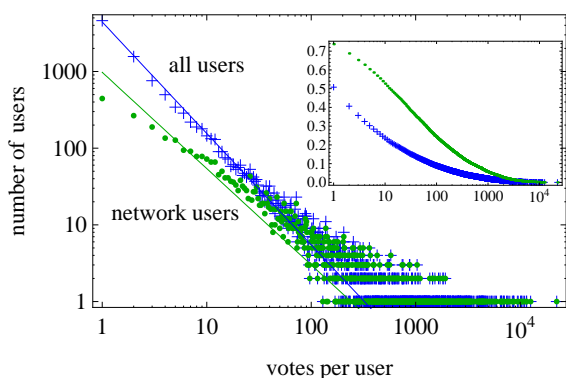


Fig. 2: Distribution of user votes. *Plus symbols*. The distribution of the number of users vs. the number of votes a user made. The solid curve indicates a power-law distribution fit to the values, with exponent  $\nu = 1.45 \pm 0.01$ . In this and other figures the range given with the parameter estimate is the 95% confidence interval. The parameter estimates and confidence intervals in this and the other figures are maximum likelihood estimates [14,15] assuming independent samples. The plot does not include the 2984 users with zero votes. *Dots*. The same distribution, but considering only users who have at least one network connection. The exponent of the indicated power-law fit is  $\nu = 1.26 \pm 0.008$ . The inset shows the cumulative version of the normalized distributions (so that they are probability distribution functions), summed up from the right.

at the same time can form links. The first factor gives more links for those who vote a lot (due to being more likely to have votes in common with others). This leads to, in effect, preferential attachment for forming links (those with more links are likely to be users with more votes, hence more overlap with others), while the attachment probability does not explicitly depend on degrees, similarly to network generation models where hidden variables determine link formation [12,13]. The activity constraint limits the link growth, corresponding to descriptive models giving truncated power-law degree distributions [9].

To make a connection between voting activities and link formation, consider Fig. 2, where we measured the number of networked users that voted on a given number of resolves. About a fifth of users have no votes on noninitial resolves. For the rest of the users, the figure shows the distribution of votes among users who voted at least once for noninitial resolves. These votes are close to a power-law distribution in the number of votes, with the number of users with  $v$  votes proportional to  $v^{-\nu}$ . Suppose that user  $A$  voted on  $N_A$  resolves, while user  $B$  voted on  $N_B$  resolves, where  $N_A$  and  $N_B$  are drawn from the probability distribution corresponding to Fig. 2 (dots). The cornerstone of the linking model we propose is the assumption that  $A$  and  $B$  form a link with a probability proportional to the product  $N_A N_B$  (justification is given below).

Caldarelli *et al.* [12] have introduced a model of network formation where vertices possess intrinsic “fitnesses”,

which are assigned to the vertices randomly, but are drawn from a probability distribution function given by a power law  $\phi(\cdot)$ . Links are then made between vertices with a probability that depends on the fitnesses of the nodes to be linked; if the functional form of the linking probability is given as the product of the two fitness values, then the resulting network will have a degree distribution that is the same as the power-law distribution associated with the node fitnesses,  $\phi(\cdot)$ .

Fitting truncated power laws to the degree distributions with exponents  $\tau$  and cutoffs  $\kappa$  of the three networks shown in Fig. 1, we find the parameters  $\tau_F = 1.25 \pm 0.04$ ,  $\kappa_F = 27 \pm 4$ ;  $\tau_A = 1.20 \pm 0.04$ ,  $\kappa_A = 59 \pm 9$ ; and  $\tau_N = 1.44 \pm 0.11$ ,  $\kappa_N = 18 \pm 6$  for the friends, allies, and nemeses networks, respectively. On the other hand, if we consider the number of votes for users who have at least one network link (Fig. 2) as playing the role of fitnesses above, we expect that a network that is formed as through the fitness model will have the same degree exponent as the exponent of the votes per user distribution. Indeed, according to our measurements this distribution follows a power law with an exponent of  $-1.26$ , closely matching the exponent of the ally ( $-1.20$ ) network, and differing slightly from that of the nemesis network ( $-1.44$ ). The truncation of the power laws seen in the degree distributions is most probably the result of the finite size of the network or of vertices gradually becoming inactive in time. While the friends network does not supposedly arise as a result of voting activity, its degree exponent  $-1.25$  is close to  $-1.26$  as well. All three networks exhibit significant positive correlation between degree and number of votes, which is largest for the allies network [7].

Ally and nemesis networks are thus most likely in close relation to voting activity. Since the network links are provided as a snapshot at the end of the data collection period only, we cannot rule out the possibility that at the time when a given link was formed the vote per user distribution looked different from what is observed at the end as is shown in Fig. 2, and thus the reasoning above would not apply for the degree distribution. For this reason we checked the votes per user distributions at four different times that were spaced equally within the whole data set comprising of 497 days. To decide whether the first three samples are identically distributed as the last one, we performed a Wilcoxon rank sum test on these [16], resulting in  $p$ -values of 0.0065, 0.29, and 0.91, in time order of the samples, respectively. Thus at the 5% confidence level, only the earliest sample can be rejected as not coming from the same distribution. This suggests that whenever a link was formed, the vote count frequency distribution could be described by the same power-law form as at the end of the data set.

A likely mechanism that we referred to above that leads to the linking probability being proportional to the product of users’ vote counts is the following: this product is naturally connected to the expected number of common resolves that two persons  $A$  and  $B$  vote on, assuming that

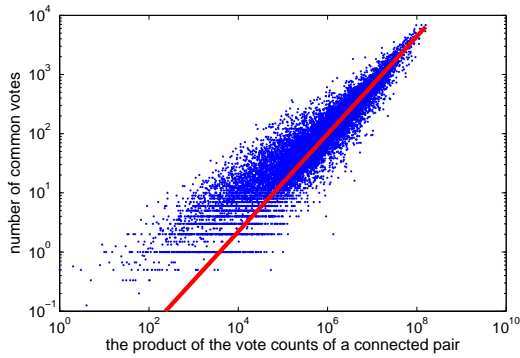


Fig. 3: The relationship between the product of vote counts for every pair of allies, and the number of their common votes. A power law has been fitted to the data indicated by the guide line, with an exponent of  $0.825 \pm 0.005$ . The prefactor to the power law is  $(1.1 \pm 0.096) \cdot 10^{-3}$ . The pairs of allies are binned by the product of their vote counts, and the overlap is averaged in the bins; the average number of common votes may thus be less than 1.

they choose the resolves randomly from the pool of all available resolves. In this case, if we have  $R$  resolves and the choices are independent of each other and uniformly distributed among the resolves, the expected number of common votes on resolves will be  $N_A N_B / R$  (the notations are as before).

To see this, with random selection the number of ways a pair can pick  $N_A$  and  $N_B$  resolves is

$$\binom{R}{N_A} \binom{R}{N_B}. \quad (1)$$

The number of ways these selections can have  $C$  resolves in common is expressed by the multinomial coefficient, accounting for the number of ways to pick the common resolves, and the remaining  $N_A - C$  and  $N_B - C$  each person picks that are distinct:

$$\frac{R!}{C!(N_A - C)!(N_B - C)!(R - N_A - N_B - C)!}. \quad (2)$$

The ratio of these quantities is the probability a pair of users, with  $N_A$  and  $N_B$  resolves, have  $C$  resolves in common. While this distribution of common votes depends separately on both  $N_A$  and  $N_B$ , the expected value of  $C$  is yielded as  $N_A N_B / R$ , and depends only on the product of the number of votes by the two users.

To check this model of common votes, we plotted the average number of common resolves as a function of the product  $N_A N_B$  for every pair  $(A, B)$  in the ally network in Fig. 3. The relationship is close to linear, insofar as the best power-law fit has an exponent of  $0.825 \pm 0.005$ . The results for the nemesis network are very similar, and a least-squares fit yields an exponent of  $0.829 \pm 0.01$  as well. We also checked the scaling for a large number of unconnected pairs of users, finding an exponent of about 0.9. It is thus plausible that voting on common resolves is a main

reason for network formation, and the deviation from a linear relationship may be due to the fact that after links are formed, the observed number of common votes for a pair  $(A, B)$  does not scale any longer as  $N_A N_B$ , since the pair sees each other’s resolves and thus their choices will be biased. However, because our data set does not contain the times when links were formed, we only know if two users were connected or not at the end of the data collection period. Thus the challenge to evaluating this mechanism is causation: resolves voted on by network neighbors are highlighted in the user interface, making them more visible and hence more likely to receive votes. On the other hand, the Essembly website makes it easy to find similar users to link to based on common votes. These two facts limit our capacity to separate voting activity that *leads* to ideological link formation and the voting on observed common resolves that is the *result* of the presence of a link. Thus common votes increase the chances of forming a link by providing information to form a profile, and links increase the chance of common votes through visibility of resolves.

Moreover, the best fit to the observed relationship shown in Fig. 3 indicates that  $c_{(A,B)} \approx 0.0011(N_A N_B)^{0.83}$ , where  $c_{(A,B)}$  is the number of common votes of link  $(A, B)$ . The prefactor 0.0011 can be interpreted as being related to the average effective size  $R_{\text{eff}}$  of the set of resolves that users choose resolves from when they vote, as elucidated by the relationship for the expected number of common votes  $c_{(A,B)} = (N_A N_B) / R_{\text{eff}}$ . While the total number of resolves is 24953 at the end of the data collection, the fact that  $R_{\text{eff}} \approx 902$  is smaller than that may be because of the following: (i) that we consider an average over the whole duration of the data set, for which the number of available resolves to vote on is not constant, but grows in time; (ii) and more importantly, that users mostly vote on recent resolves only, virtually decreasing the scope of possible resolves that may result in an overlap between a pair’s choices.

Fig. 4 shows that the types of links vary depending on user activity. For this plot, users with at least one network connection are grouped into quantiles by their number of votes. Each point on the plot is the average fraction of link types among users in that quantile, with the error bar indicating the standard deviation of this estimate of the mean. Users with few votes tend to have most of their links to friends only, so do not participate much in the ideological networks. On the other hand, users with many votes tend to have most of their links in the ideological networks and to people who are not also friends. One way to quantify the trends shown in the figure is through the correlation coefficient between number of votes and fraction of each link type of all users with at least one network connection:  $-0.24 \pm 0.02$ ,  $-0.003 \pm 0.01$  and  $0.27 \pm 0.03$  for “only friends”, “friends & ideological” and “non-friends” link types, respectively, with the ranges showing the 95% confidence interval of the correlation from a bootstrap test [17]. The signs of these correlations correspond to the trends

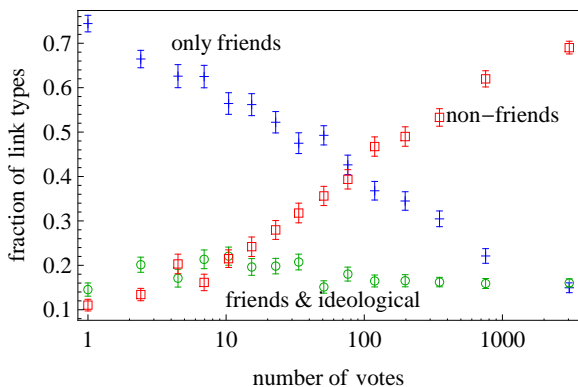


Fig. 4: Fraction of link types vs. number of votes on noninitial resolves. A linked pair of users are denoted “only friends” when their only link is in the friends network, “non-friends” when they are not linked in the friends network, and “friends & ideological” when they have a friends link as well as a link in the allies or nemeses networks.

seen in Fig. 4, with the fraction of “friends & ideological” links consistent with no correlation with number of votes. The same trend in link types occurs as a function of other measures of user activity, such as the time a user is active or the number of links a user has.

**Users.** — In this section we are going to study the individual users’ activities. Fig. 5 shows most users are active for only a short time (less than a day), as measured by the time between their first and last votes (this includes votes on the initial resolves during registration—users need not vote on all of them immediately). The 4762 users active for at least one day account for most of the votes and links, and we focus on these *active users* for our model. For these users, Fig. 5 shows an exponential fit to the activity distribution for intermediate times. Thus users who have sufficient interest in the system to participate for at least a few days behave approximately as if they decide to stop participating as a Poisson process. The additional decrease at long times (above 200 days) is due to the finite length of our data sample (about 500 days).

The distribution of number votes per user arises from two factors: how long users participate before becoming inactive, and how often they vote while active.

New users arrive in the system when they register, and we model this as a Poisson process with rate  $\alpha$ , and users leave the system (*i.e.*, become inactive) with a rate  $1/\tau$ .

User activity is clumped in time, with groups of many votes close in time separated by gaps of at least several hours. This temporal structure can be viewed as a sequence of user sessions. We measured the averaged distributions for interevent times between activities of individuals, which show long-tail behavior, similarly to other observed human activity patterns, such as email communications or web site visits [18]. To model the number of votes per user in the long time limit where we are only interested in the total number of accumulated votes for a

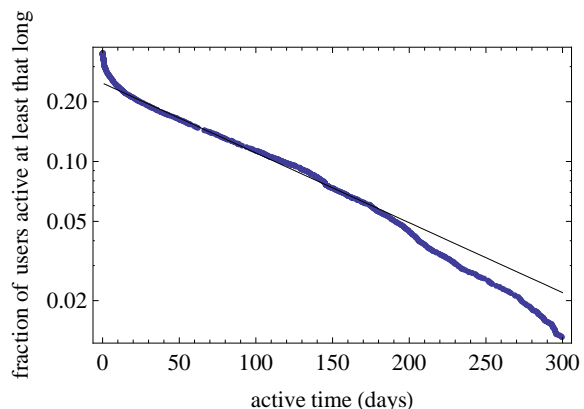


Fig. 5: Distribution of activity times for users. The line shows an exponential fit to the values between 10 and 200 days, proportional to  $e^{-t/\tau}$  where  $\tau = 124$  days.

particular user, this clumping of votes in time is not important. Specifically we suppose each user has an average activity rate  $\rho$  while they are active on the site (cf. Fig. 5), given as  $\rho_U = e_U/T_U$ , where  $\rho_U$  is user  $U$ ’s activity,  $e_U$  is his/her number of events (*i.e.*, votes, resolve creations, and links), and  $T_U$  is the time elapsed between her first and last vote. We suppose the  $\rho_U$  values arise as independent choices from a distribution  $P_{\text{user}}(\rho_U)$  and the values are independent of the length of time a user is active on the site. These properties are only weakly correlated (correlation coefficient  $-0.06$  among active users).

We characterize user activities by fractions  $q$  and  $\lambda$  for creating resolves and forming links, respectively. The rate of voting on existing resolves for a user is then  $\rho_U(1 - q - \lambda)$ , which is by far the most common of the three user activities. For simplicity, we treat these choices as independent, and take  $q$  and  $\lambda$  to be the same for all users. Thus in our model, the variation among users is due to their differing overall activity rates  $\rho_U$  and amount of time they are active on the site  $T_U$ .

We estimate the model parameters from the observed user activities, and restrict attention to active users. Table 1 shows the estimates for the parameters, governing participation ( $\alpha$  and  $\tau$ ) and activity ( $q$  and  $\lambda$ ) choices. Fig. 6 shows the observed cumulative distribution  $\rho_U$  values and a fit to a lognormal distribution.

The distributions of activity times and rates presumably reflect the range of dedication of users to the site, where most users are trying the service for a very limited time, but active users are also represented in the heavy tail. Such extended distributions of user activity rates is also seen in other activities, including use of web sites such as Digg [5] and scientific productivity [19]. We treat the distribution of user participation as exogenous, but it could also depend on a user’s experience with the site [20].

Finally, our model also describes the significant fraction of users who form no links as due to a combination of low activity rate  $\rho$  and short activity time  $t$ . Specifically,

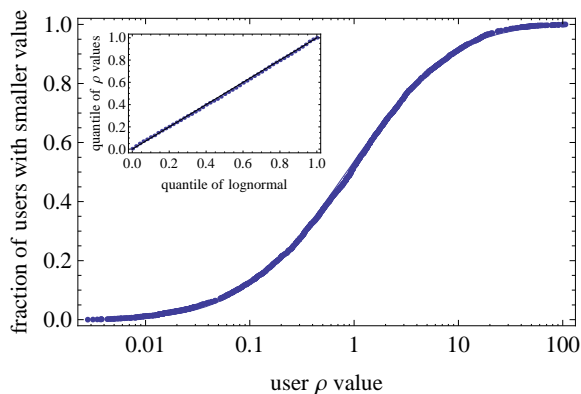


Fig. 6: Cumulative probability distribution function of activity rates,  $\rho_U$ , for the 4741 users who were active at least one day and voted on at least one noninitial resolve or formed at least one link. Plot includes a curve for a lognormal distribution fit, which is indistinguishable from the points and with parameters  $\mu = 0.03 \pm 0.05$  and  $\sigma = 1.70 \pm 0.04$ . The  $\rho$  values are in units of actions per day. The inset shows a quantile-quantile plot comparing the observed  $\rho_U$  values with the fitted lognormal distribution, which displays an excellent match.

in our model the probability that a user has no links is  $e^{-\lambda\rho t}$ . For active users, whose activity time distribution is roughly exponential with time constant  $\tau$ , the values in Table 1 and the distribution of  $\rho$  values in Fig. 6 give the probability for no links as the average value of  $e^{-\lambda\rho t}$  equal to 23%. This compares with the 1242 out of 4762 active users (26%) who have no links in our data set.

**Discussion.** — We centered our investigations around two areas: how online social networks form around topical interests, and the wide range in user activity levels in online service participation.

In particular, first, we gave a plausible, quantitative explanation of the long-tailed degree distributions observed in online communities, based on only the activity patterns of users and the underlying collaborative mechanisms. This involved showing that peer selection to form ideological relationships is explicable by the scope of shared interests, and moreover that the interface through which potential contacts learn about each other may play a crucial part in this. Therefore the need to build ideological profiles suggests votes on common resolves is key to the number and type of links. Also, a user forming a link is more likely to have many common votes with other users who are very active (and hence have many votes). Thus forming links based on common votes is likely to lead people to link preferentially to highly active users, who will in turn tend to have many links. We also found, second, that most users try the online services only briefly, so most of the activity arises from a relatively small fraction of users who account for the diverse behavior observed.

The implications may extend beyond the scope of purely online societies to describe other societal connections as

parameter	value
new user rate	$\alpha = 9.3/\text{day}$
activity time constant	$\tau = 124$ days
resolve creation	$q = 0.018 \pm 0.0002$
link creation	$\lambda = 0.043 \pm 0.0003$

Table 1: User activity parameters.

well where shared interests motivate relationship formation. Our model, however, does not address other significant properties of the networks, such as community structure and assortativity, and why they differ among the three networks; for a discussion of some of these points, see Ref. [7]. Nor does our model address detailed effects on user behavior due to their network neighbors.

## REFERENCES

- [1] LAM C., in *Proc. of the 5th ACM Conf. on Electronic Commerce (EC'04)* (ACM Press) 2004 pp. 254–255.
- [2] GUHA R., KUMAR R., RAGHAVAN P. and TOMKINS A., in *Proc. of the 13th Intl. World Wide Web Conf. (WWW2004)* (ACM, New York) 2004 pp. 403–412.
- [3] MITZENMACHER M., *Internet Mathematics*, **1** (2004) 226.
- [4] SORNETTE D., *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools* 2nd Edition (Springer, Berlin) 2004.
- [5] LERMAN K., in *IEEE/WIC/ACM Intl. Conf. on Web Intelligence and Intelligent Agent Technology* 2007 pp. 255–258.
- [6] ANDERSON C., *The Long Tail: Why the Future of Business is Selling Less of More* (Hyperion) 2006.
- [7] HOGG T., WILKINSON D. M., SZABO G. and BRZOWSKI M., in *Proc. of the AAAI Symposium on Social Information Processing* 2008 pp. 30–35.
- [8] NEWMAN M. E. J., STROGATZ S. H. and WATTS D. J., *Phys. Rev. E*, **64** (2001) 026118.
- [9] AMARAL L. A. N., SCALA A., BARTHELEMY M. and STANLEY H. E., *Proc. of the Natl. Acad. Sci.*, **97** (2000) 11149.
- [10] VÁZQUEZ A., *Phys. Rev. E*, **67** (2003) 056104.
- [11] DOROGOVTSSEV S. N. and MENDES J. F. F., *Adv. Phys.*, **51** (2002) 1079.
- [12] CALDARELLI G., CAPOCCI A., DE LOS RIOS P. and MUNOZ M. A., *Phys. Rev. Lett.*, **89** (2002) 258702.
- [13] SERRANO M. A., KRIOUKOV D. and BOGUNA M., *Phys. Rev. Lett.*, **100** (2008) 078701.
- [14] NEWMAN M. E. J., *Contemporary Physics*, **46** (2005) 323.
- [15] JAMES A. and PLANK M. J., arxiv.org preprint 0712.0613 (2007).
- [16] WILCOXON F., *Biometrics*, **1** (1945) 80.
- [17] EFRON B., *Annals of Statistics*, **7** (1979) 1.
- [18] VÁZQUEZ A., OLIVEIRA J. G., DEZSO Z., GOH K.-I., KONDOR I. and BARABASI A.-L., *Phys. Rev. E*, **73** (2006) 036127.
- [19] SHOCKLEY W., *Proc. of the IRE*, **45** (1957) 279.
- [20] WILKINSON D. M., in *Proc. of the 2008 ACM Conf. on E-Commerce* 2008 pp. 302–309.