Friends and Foes: Ideological social networking

Michael J. Brzozowski, Tad Hogg, Gabor Szabo
Hewlett-Packard Labs
1501 Page Mill Road, MS 1139, Palo Alto, CA 94304, USA
{mike.brzozowski, tad.hogg, gabors}@hp.com

ABSTRACT
Traditional online social network sites use a single monolithic "friends" relationship to link users. However, users may have more in common with strangers, suggesting the use of a "similarity network" to recommend content. This paper examines the usefulness of this distinction in propagating new content. Using both macroscopic and microscopic social dynamics, we present an analysis of Essembly, an ideological social network that semantically distinguishes between friends and ideological *allies* and *nemeses*. Although users have greater similarity with their allies than their friends and nemeses, surprisingly, the allies network does not affect voting behavior, despite being as large as the friends network. In contrast, users are influenced differently by their friends and nemeses, indicating that people use these networks for distinct purposes. We suggest resulting design implications for social content aggregation services and recommender systems.

Author Keywords
Social networks, online communities, Essembly.

ACM Classification Keywords
H5.3. Information interfaces and presentation (*e.g.*, HCI): Group and organization interfaces. J4. Social and behavioral sciences.

INTRODUCTION

Traditionally, online social network sites like Facebook(.com) and MySpace(.com) allow people to form links to "friends" but do little to qualify the semantic meaning of the friendship (although Facebook permits users to annotate how they met a friend). As a result, many users "collect" friends on these sites, conflating "acquaintances" with "friends" [2]. Since there's often a social stigma against rejecting friendship offers [3], many users' friends expand to include people they don't really know. It's usually in users' strategic interests to collect these "weak ties", which may become useful in spreading or seeking information [7].

In designing social Internet services, it's common to add a networking component to allow people to influence each others' discovery of relevant content and products. This collaborative filtering may be done implicitly, as on amazon.com or last.fm, or explicitly, as on Facebook and del.icio.us. Social "news aggregators" like Digg(.com) even rely on this behavior—stories promoted to the front page usually must spread through a network of friends to gain popularity [8]. Recommender systems based on collaborative filtering effectively automate this process by suggesting items popular with people with similar interests, who are not necessarily friends [6, 10]. In general, however, recommendations are trusted more when they appear to come from trusted humans than computer algorithms [11].

For sites targeting a diverse set of interests, attracting a well-connected community is key to stimulating user interest in content. We hypothesize that improving the semantic granularity of "friendship" in a social network will increase the relevance of friends' influences in filtering content.

Essembly

We examined user behavior from Essembly(.com), a "fiercely non-partisan social network" that allows members to post resolves reflecting controversial opinions, *e.g.* "Overall, free trade is good for American workers." Members can then vote on these resolves, using a four-point scale: *Agree*, *Lean Agree*, *Lean Against*, or *Against*. Users can only vote once per resolve, and all their votes are viewable by other members, forming an ideological profile.

About 44% of all votes recorded were *Agree* votes, and about 34% of votes were *Against*, indicating that users typically feel strongly about the resolves they vote on. However, 56% of all votes were *Agree* or *Lean Agree*, perhaps suggesting that people either tend to submit resolves agreeable to the mainstream Essembly community or that people feel bad about voting *Against* resolves.

The site differs from MySpace, which offers only generic "friends" connections (supposedly representing a real-life social network). Essembly is unique in that it defines three semantically distinct but overlapping types of connections:

- *Friend:* "someone you know personally and have a friendship with in the real world. You have a personal, not just professional, relationship with this person."

CHI 2008, April 5–10, 2008, Florence, Italy.

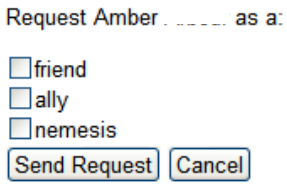

Figure 1. Essembly defines three connections.

- *Ally:* "someone who you don't necessarily know, but…share a desire to make some change in the world."
- *Nemesis:* "someone who you don't agree with…their world view is just psychotically skewed."

Essembly acknowledges that friends may be allies or nemeses as well, and enables people to request multiple connections at the same time (see Figure 1). Indeed, 29% of friendships are alliances, and 25% of alliances are between friends. These links are explicitly created by users, and must be approved by both parties (they are bidirectional).

The friends and allies networks are similar in size. 4,873 users have 13,516 friendships, and 3,228 users participate in 15,634 alliances; while more people have friends, the allies network is denser. The nemeses network is relatively smaller: 1,117 people created 1,953 nemesis relationships.

While all resolves are accessible by browsing and search, the site encourages users to vote on resolves people in their networks have voted on. The home page shows how their friends, allies, and nemeses voted on recently active resolves (see Figure 2). This enables rich ideological comparisons between connected people; calibrating these comparisons provides additional incentive to vote on resolves in common. These networks facilitate finding interesting resolves for users, as a sort of "social filter".

Before voting, users can see how individual users voted, but a resolve's aggregate vote counts are not exposed until after a user votes. Users often add comments as a way of explaining or justifying their vote. For instance, one user among a 15% minority who voted against the resolve "Hummers are ugly and ridiculous cars" qualified his vote: "I like the newer models…The older hummers *[sic]* were eyesores."

Our Dataset

Our dataset consists of fully anonymized records, representing a complete history of voting activity on the site from August 2005 through December 12, 2006, and the final state of network connections (*i.e.*, friends, allies, and nemeses). The data includes 1.4 million votes from 15,092 unique registered users with the time of each vote. We do not have the times when users created links. The median number of votes per user is 11, but there is a large variance (average μ = 93; standard deviation σ = 529). Of these users, 5,027 declared friends, allies, and/or nemeses.

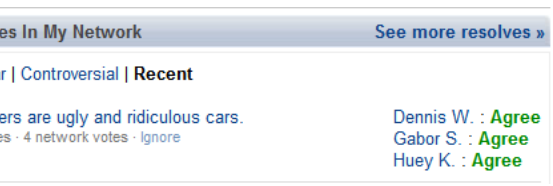

Figure 2. Active resolves in a user's network are shown on the home page.

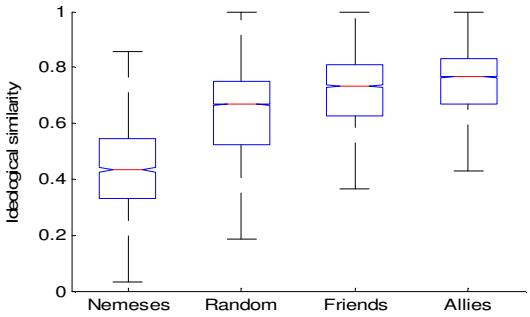

Figure 3. Ideological similarity between pairs of people.

We analyzed 24,963 resolves, most of which received 40 or more votes (μ = 57; σ = 238). Ten standard initial resolves are presented to new users when they first sign up; while not mandatory, each one received votes from up to 82% of users. To protect user anonymity, we do not have access to the content of resolves or comments.

NETWORK HOMOPHILY

In social networks, *homophily* is the principle that people tend to be connected with people demographically and behaviorally similar to them. This structures all sorts of network ties, including marriage and friendship in the "real world" [9]. It's also reflected in online dating [5] and social networks [1].

Essembly provides a metric of "ideological similarity", defined for users *a* and *b* as:

$$S(a,b) = 1 - \underset{r \in R(a,b)}{mean}\left(\frac{|V(a,r) - V(b,r)|}{3}\right)$$

where $V(a, r)$ is user a's vote on resolve r on a 1-4 scale (1 = *Agree*, 4 = *Against*), and $R(a, b)$ is the set of all resolves a and b have both voted on. Therefore two users who are in complete agreement have similarity 1, while two users who vote in complete opposition (*i.e.*, one votes *Agree* while the other votes *Against*) would have similarity 0. The median ideological similarity between a pair of random users is 0.67. This is partially because the vote distribution is somewhat uneven; the average vote is 2.3, so it's difficult to be *too* dissimilar.

As shown in the box-and-whisker plot in Figure 3, the three networks have fairly different distributions from random pairings of people. Friends and allies tend to be more ideologically similar (medians 0.73 and 0.77, respectively), while nemeses are decidedly *dissimilar* (median 0.43). In short, people are more likely to agree with their allies than their friends, and with their friends than their nemeses. This provides strong evidence that the three semantically distinct networks capture different levels of ideological homophily.

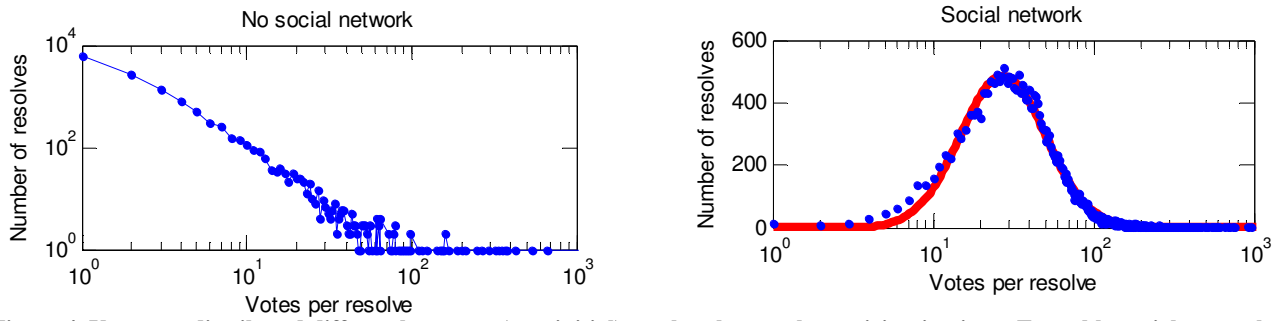

Figure 4. Votes are distributed differently among (non-initial) resolves by people participating in an Essembly social network.

SOCIAL INFLUENCES

Users who participate in one or more Essembly social networks tend to vote on more resolves. Users without social connections voted on a median of 11 resolves each, while users with friends, allies, and/or nemeses had a median of 26 votes.

People use their social networks to discover resolves. Users without friends, allies, or nemeses gravitate towards more popular resolves, distributing votes across resolves in a power law. This is a classic "rich-get-richer" mechanism, by which more popular resolves attract proportionally more votes (see Figure 4). By contrast, users who have social networks on Essembly distribute their votes differently: a lognormal distribution with $\mu = 3.7$, $\sigma = 0.6$. This difference suggests that individual participation in the social network correlates with what resolves a user chooses to vote on.

There is homophily in aggregate voting patterns—both in how friends vote and what they vote on. As a more detailed question, we examined whether social influences affect voting on the micro level. To do this, we considered 1.3 million votes by users during the study period, excluding votes on the ten initial resolves, which are presented to everyone on signup. To build a predictive model, we also needed examples of cases where users *could have* voted on something but didn't. But it's possible, for example, for a user to be presented with three active resolves (see Figure 2) and vote on them within minutes of each other. Our data doesn't specify the order resolves were presented to the user. So voting on one resolve first does not necessarily indicate a lack of interest in the other two.

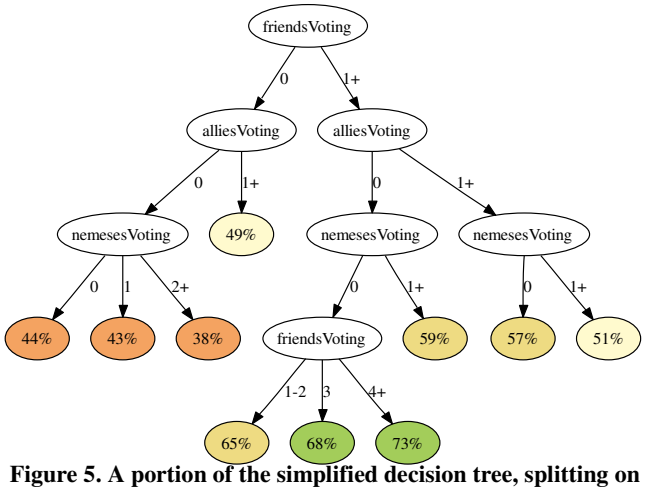

Figure 5. A portion of the simplified decision tree, splitting on the number of friends, allies, and nemeses voting previously.

We solved this problem by inferring user *sessions*: periods of activity for a given user, punctuated by three or more hours without any votes. At the end of each session, for each vote a user placed during the session, we created one negative example by randomly selecting one of the resolves the user *could have* voted on (from the set of all resolves that existed at the time) but didn't. This results in a roughly equal number of positive and negative examples for our learning algorithms (*prior* P(vote) = 0.5). So if we choose some subset of samples that is also 50% positive (*posterior* P(vote) = 50%), the criteria used to select the sample probably don't affect voting behavior.

Decision Tree

We learned a decision tree over the complete set of examples using [4]. The leaves show the posterior probability a user will vote on a resolve under the given conditions, relative to the prior distribution. As features we used the absolute number of friends, allies, and nemeses who'd voted previously and thus could have shown up on a user's home page as an influence. A portion of the tree is shown in Figure 5.

Rows 1 and 2 of Table 1 suggest users are about 48% (i.e., 65 / 44 – 1 = 0.48) more likely to vote on a resolve if one or more of their friends voted on it than if none of their neighbors did ($p < 0.001$). By contrast, if four or more friends (and no one else) voted on a resolve already, they are 66% more likely to vote on it (see Table 1 rows 1, 3).

Curiously, we observed that in cases where two or more nemeses voted on a resolve (but no friends or allies), the posterior P(vote) drops to 38%. That is, a user is 14% *less*

Friends	Allies	Nemeses	*Post.* P(vote)	*N*
0	0	0	44%	804,299
1+	0	0	65%	96,910
4+	0	0	73%	3,022
0	0	2+	38%	12,462
0	1+	0	49%	604,042
0	1+		49%	825,656

Table 1. Posterior probability of voting, as a function of how many friends, allies, and nemeses voted previously.

likely to vote on a resolve if two or more nemeses voted on it than if no one in his or her local networks did ($p < 0.001$; see Table 1 rows 1, 4). This is somewhat unexpected: why bother confirming someone as a nemesis, thereby making their votes visible to you as part of your network, if you're just going to ignore resolves they vote on?

It is possible that nemeses tend to vote on resolves whose premises are so "psychotically skewed" that they're perceived as ludicrous and not worth considering, effectively using nemeses to filter out their resolves. Another possibility may be that users feel disinclined to vote on resolves when they know they'll disagree; this may also explain the slight bias toward *Agree* votes described earlier.

Although Essembly encourages users to vote on resolves voted on by their allies, and there are indeed many examples where a user sees only allies voting on a resolve, this does not appear to be happening. Users are not significantly more likely to vote on resolves their allies have voted on (see Table 1 row 5).

Users are also 33% more likely to vote on a resolve that friends voted on than some combination of allies and nemeses ($p < 0.001$; see Table 1 rows 2, 6). This suggests that friends are more influential than allies or nemeses in selecting resolves.

Many other nodes are more ambiguous; we suspect this is because at most three people in a user's network are shown as possible influences on the home page (see Figure 2). Other noise arises from alternate ways a user can discover resolves (*e.g.*, recent or globally popular resolves).

CONCLUSION AND FUTURE WORK

We demonstrated, through a combination of microscopic and macroscopic dynamics, that the presence of a social network affects *what* users will vote on and *how* they will vote. Refining the granularity of friendship semantics increases the meaning of online "friendship", and likewise the influence these friends have on others' behavior.

This has implications for the design of social networking and recommender services. Sites that encourage online socialization and registration of users' "friends" may fall victim to the dilution of "friendship" as seen on Friendster [3]. As the network grows, people may be compelled to add as "friends" people whom they know only superficially or who don't hold much influence over them. The use of this friends network to make recommendations or to influence users may then be diminished.

Therefore, sites seeking to exploit the *influence* exerted by friends in changing behavior or making recommendations should provide multiple classes of friendship, or some way to ignore non-influential friends. Recommendations are more persuasive coming from real friends than "similar" people, whether generated by collaborative filtering [11] or by explicitly chosen allies. The closer a user is to recommending friends, the stronger the persuasion is likely to be.

In addition, nemeses or blacklists are a good way to filter out uninteresting content. While allies may not influence behavior, they offer an alternative type of connection, possibly preserving the meaning of friendship. These extra connections allow people to distinguish interesting or uninteresting people, distinct from the emotionally loaded notion of friendship.

Future study is needed to better isolate this effect, perhaps in a more controlled environment. It may also prove fruitful to consider alternate models of user behavior, and to draw parallels with similar data from other social network services. It also remains to be seen to what extent this practice encourages people to make their declared online friends resemble their real world friends.

ACKNOWLEDGMENTS

We gratefully thank Chris Chan and Jimmy Kittiyachavalit of Essembly for providing access to this data. We also thank Dennis Wilkinson, Scott Golder, and Kostas Bimpikis for helpful conversation.

REFERENCES

1. Adamic, L. A., Buyukkokten, O., and Adar, E. A social network caught in the Web. *First Monday, 8*, 6 (2003).
2. boyd, d. Friends, friendsters, and top 8: Writing community into being on social network sites. *First Monday, 11*, 12 (2006).
3. boyd, d. m. Friendster and publicly articulated social networking. In *Ext. Abstracts CHI 2004*, ACM Press, 1279-1282.
4. Chickering, D. M. The WinMine Toolkit. Microsoft, Redmond, WA. MSR-TR-2002-103.
5. Fiore, A. T., and Donath, J. S. Homophily in online dating: when do you like someone like yourself? In *Ext. Abstracts CHI 2005*, ACM Press, 1371-1374.
6. Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. Using collaborative filtering to weave an information tapestry. *Commun. ACM, 35*, 12 (1992), 61-70.
7. Granovetter, M. The strength of weak ties: A network theory revisited. *Sociological Theory 1* (1983), 201-233.
8. Lerman, K. Social Networks and Social Information Filtering on Digg. In *Proc. Int. Conf. on Weblogs and Social Media* (2007).
9. McPherson, M., Smith-Lovin, L., and Cook, J. M. Birds of a feather: Homophily in Social Networks. *Annual Review of Sociology, 27,* 1 (2001), 415-444.
10. Shardanand, U., and Maes, P. Social information filtering: algorithms for automating "word of mouth". In *Proc. CHI 1995*, ACM Press, 210-217.
11. Sinha, R. R., and Swearingen, K. Comparing recommendations made by online systems and friends. In *Proc. DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries* (2001)