

You Are Who You Talk To: Detecting Roles in Usenet Newsgroups

Danyel Fisher, Marc Smith, and Howard T. Welser

Microsoft Research

danyelf@microsoft.com, masmith@microsoft.com, twelser@u.washington.edu

Abstract

Understanding the social roles of the members a group can help to understand the social context of the group. We present a method of applying social network analysis to support the task of characterizing authors in Usenet newsgroups. We compute and visualize networks created by patterns of replies for each author in selected newsgroups and find that second-degree ego-centric networks give us clear distinctions between different types of authors and newsgroups.

Results show that newsgroups vary in terms of the populations of participants and the roles that they play. Newsgroups can be characterized by populations that include question and answer newsgroups, conversational newsgroups, social support newsgroups, and flame newsgroups.

This approach has applications for both researchers seeking to characterize different types of social cyberspaces as well as participants seeking to distinguish interaction partners and content authors.

1. Introduction

Joining a new social group can be a disorienting experience. Understanding the social roles of different members of a group is a good way to understand both the social context of the group, and information about the people within it. A number of different visualization tools exist [2, 10, 12] that are meant to give some sense of how a conversation space is constructed. These various visualizations have all shown different perspectives on metadata within online social spaces [11].

In this paper we present a way of applying social network techniques to online conversation spaces. In particular, we focus on the question “who is this person with whom I am interacting?” looking to

understand how their social behavior suggests their role within the space. By generating these networks for each author in a newsgroup, we can begin to roughly characterize the newsgroup in terms of the different types of participant authors.

In this paper, we examine the social networks created by patterns of reply between authors contributing messages to nine Usenet newsgroups during a one-month period (Table 1). These nine groups were hand-selected to provide a broad overview of Usenet topics. These newsgroups are of medium size, relative to the rest of Usenet: they are not among either the biggest groups, or the smallest.

We sought out groups that we already knew were of several general genres: question and answer, social support, discussion, and flame. We chose these newsgroups because they are exemplars of their types. *Microsoft.public.windows.server.general*, for example, is low on discussion, but active with questions and answers; in contrast, *alt.flame* is uncluttered by rational discussion.

There have been a number of studies that have attempted to examine and describe the sets of social roles that occur in both online and offline conversation [6, 11]. In this paper, we expand on the ideas of Turner et al [11] and Welser [14], in examining a small number of visibly different roles.

A social role “refers to the behavior of status-occupants that is oriented toward the patterned expectations of others” (pg 41) [7]. This notion emphasizes structure – in terms of relations to others, and structure, in terms of meaningful expectations for systematic behavior. We focus on emergent roles that grow from participation in the focal activities, as distinct from formal organizational roles like “manager”.

Table 1. Nine newsgroups, with some basic statistics.

Name <i>January 2001 except as noted</i>	authors	replies	% threads with		% authors with	
			2 messages	5+ messages	degree > 10	1 message
Technical Newsgroups						
comp.soft-sys.matlab	437	712	28%	20%	3%	41%
microsoft.public.window s.server.general, <i>Jan.</i> <i>2004</i>	855	1489	33%	10%	3%	51%
Discussion Newsgroups						
rec.kites	276	924	19%	32%	12%	32%
rec.music.phish	1263	6105	23%	22%	20%	29%
Political Newsgroup						
alt.politics	2187	16059	17%	40%	20%	26%
Flame Newsgroups						
alt.flame	618	3802	16%	34%	13%	32%
alt.alien- vampire.flonk.flonk.flonk	755	6494	17%	50%	16%	33%
Social Support Newsgroups						
alt.support.divorce	339	2937	14%	53%	21%	23%
alt.support.autism	188	3095	17%	60%	32%	20%

We further focus our attention on the relationship between the communication behaviors of participants in Usenet with their social structural position derived from network analysis. In network analysis, we can identify roles as positions that have a distinct pattern of relations to other positions [13]. Our analysis takes fundamental features of an actor’s structural position (like their number of neighbors) and compares how those basic attributes vary across different conversational spaces.

2. Methods

Our approach focuses on linking two scales of analysis between individual and collective structures in Usenet newsgroups. We alternately examine newsgroup behavior and the behavior of individual authors. Each casts light upon the other. The key social structure used throughout is *egocentric social networks constructed through patterns of reply*. An egocentric network reports the other people and all ties surrounding an actor to a limited distance. In this case, we have limited our distance to two: that is, we are looking at the neighbors of the neighbors of the actors.

Phrased less formally, we are trying to understand whether the people with whom the actor interacts are themselves well-connected or poorly-connected, and so we look at their neighbors.

The choice of using egocentric measures is discussed more thoroughly in [6]. It is driven by the particular properties of network structure in public

persistent conversation spaces. All messages we examine are publicly visible; thus, notions of information transfer within the group are less confining than they might be in an organizational network. In short, we expect people to choose their correspondents, but not to be thoroughly-embedded within a broader network.

Data to construct the networks presented here was generated by the Netscan system [8] at Microsoft Research. Netscan stores message headers from Usenet newsgroups and constructs a range of measures of author, newsgroup and thread size and activity. Netscan was used to generate reports showing which authors sent messages in response to messages from another author. This data was further analyzed and graphically presented making extensive use of the JUNG [9] system for visualizing and computing social network data.

Table 1 shows the nine newsgroups selected, and some overview data related to these newsgroups.

We collected data from the “conversational subset” of the newsgroup’s messages: the set of messages that are either *responded to* by someone else, or are themselves *responses* to someone else. (These do not, then, include messages that are do not receive replies.) This is examining only *replies within the month* of January, 2001. (For the more recent newsgroup *microsoft.public.windows.server.general*, we examine January of 2004).

2.1. A Person's Social Network

The core of our data is composed of social networks built from the interactions present in a set of newsgroups. This network will, as with all social networks, connect pairs of vertices with edges. Vertices represent individual authors within the newsgroup who have replied to at least one message, or been replied to at least once. Directed edges represent the replies: an edge from A to B means that A replied to B's message. Edges are further annotated with a weight, which indicates the number of times which A has replied to B. Note that we use the word "reply" slightly loosely; within newsgroups, one does not directly reply to someone else, but rather posts a message in response to another's message. The distinction is subtle, but important: while we discuss interpersonal relations in terms of who replied to whom, we acknowledge that it is possible to post a follow-up message to a thread without giving any thought to the person who is being responded to.

None the less, we will—for simplicity—elide "A posted a message, in direct response to a previous message posted by B" to "A replies to B." As this is interpreted as an edge in the graph, we will also speak of the graph properties as social interaction. Thus, "A is of high out-degree" means "A responded to many different people," while "B is of high in-degree" means "B was responded to by many people."

We visualize these networks by showing the second-degree egocentric network around any given actors: that is, we look at all of that actor's neighbors, and their neighbors, as in Figure 1b. Directed ties are depicted with arrowheads, while the weight of edges is illustrated with thickness.

We generated one- and two-degree egocentric networks for all 6918 authors in our data set, and our analysis summarizes insights from comparing these networks across the nine groups. In addition, to look at individual connectivity, we calculate a distribution of the out-degree of each actor's out-neighbors: that is, we look at whether the actor typically replies to well-connected or poorly-connected alters. We examine the *interconnectivity* of members of newsgroups, and we look at the *reciprocity of ties* within these limited networks.

2.2. Group-Wide Comparisons

We have applied several techniques to compare groups to each other. We compare the percentage of authors who have appeared only once in the group, and look at thread length. We analyze collective in-degree and out-degree statistics, and we compare degree distribution coefficients across the nine

groups. We briefly describe these three strategies here.

A basic way to contrast different newsgroups is to compare the percentage of authors who have appeared in only one message during the time period. In all newsgroups, appearing only once can be a result of many different things: some authors, for example, will post infrequently, and so only be seen once. Some authors are only visiting the newsgroup, or are unintentionally cross-posted into the newsgroup. However, differences in newsgroup behavior can also cause different reasons to appear only once. In particular, we might believe that the more inclusive a newsgroup is, the more likely it is that authors will appear more than once. We can similarly examine thread length: newsgroups that have short threads are likely to be oriented toward question and answer, while longer threads suggest discussion.

Another way to compare newsgroups is to compare their in-degree distribution and out-degree distribution. Any large newsgroup's participation level reflects a rough power-law curve: there will be a few very active participants, and a great many inactive participants. While the shape is the same, we can compare the rate at which these power laws drop off—that is, the degree of inequality in the group—between different groups. We might expect a group with a strong central core to have a very different distribution from a group that has broad participation. In addition, we can compare the distribution of in-degree to the distribution of out-degree for any given group. Intuitively, this is the process of seeing whether the most active participants are as replied-to as they do replying: are they central as people who start conversations, or as ones who end them?

Last, we can adapt the technique used with the individual network visualizations of looking at the degree distribution of out-neighbors by combining the distributions across the groups. By combining the distributions across the most active actors, we generate distributions which can be, again, compared.

3. Results

We discuss first the individual measures that identify personal roles. For each newsgroup and each role, we present a typical person who exemplifies that role, with a discussion of that sort of person and some of the characteristic identifiers for locating them within the group.

Next, we discuss group-wide statistics, and connect the group types we had previously identified to these broader statistics.

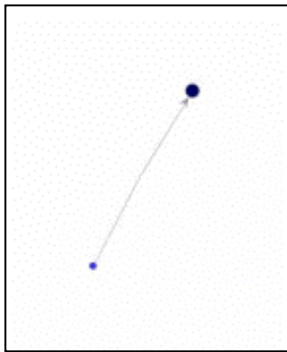
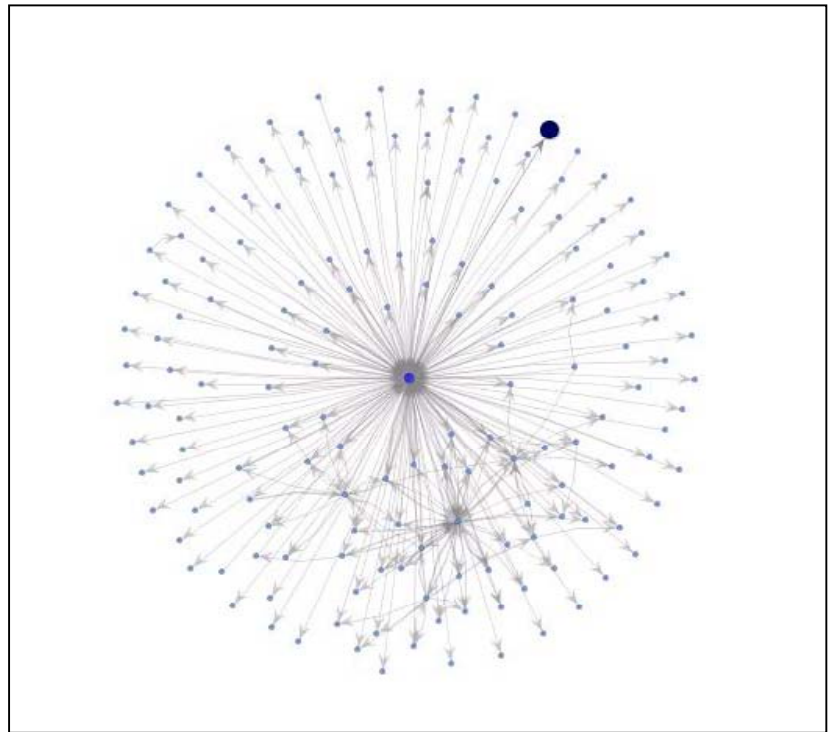


Figure 1: A question-and-answer, from the questioner's perspective, at distance 1 (above) and 2 (right).

The questioner is the large black dot; distance 1 is in blue, and distance 2 in gray.



3.1. Network Diagrams

We can refine our understanding of the differences between these newsgroups—and the individual authors within them—by examining visualizations of the network structures themselves. Let us begin with a simple example: that of a typical question and answer (Figure 1). From the perspective of the questioner, this is a simple transaction: they (the large dark dot) ask a question, and get a response. They get only one response, and never connect to any other person. We gain a greater understanding of the social roles involved, when we examine the network at **distance two**. This allows us to see the network around the person who responded.

We see that this response was not an unusual thing for this answer person. Indeed, this person seems to

send most of his or her replies to people who are disconnected from each other. There is a small cluster in the bottom-right where several people of high degree are inter-connected, but that is dwarfed by the number of individual messages.

To re-examine this from an answer person's perspective, we see an answer person in Figure 2a. Note, again, the large number of connections to nodes that are not themselves interconnected.

When we expand to distance 2, we see that some few of the people to which this person interacted were answer people – but the vast majority was not. To quantify this, we can examine the *out-degree distribution* of his neighbors: that is what was the out-degree of the people to whom this person responded? This question allows us to understand whether they preferentially reply to people who are well-connected, or people who are more isolated.

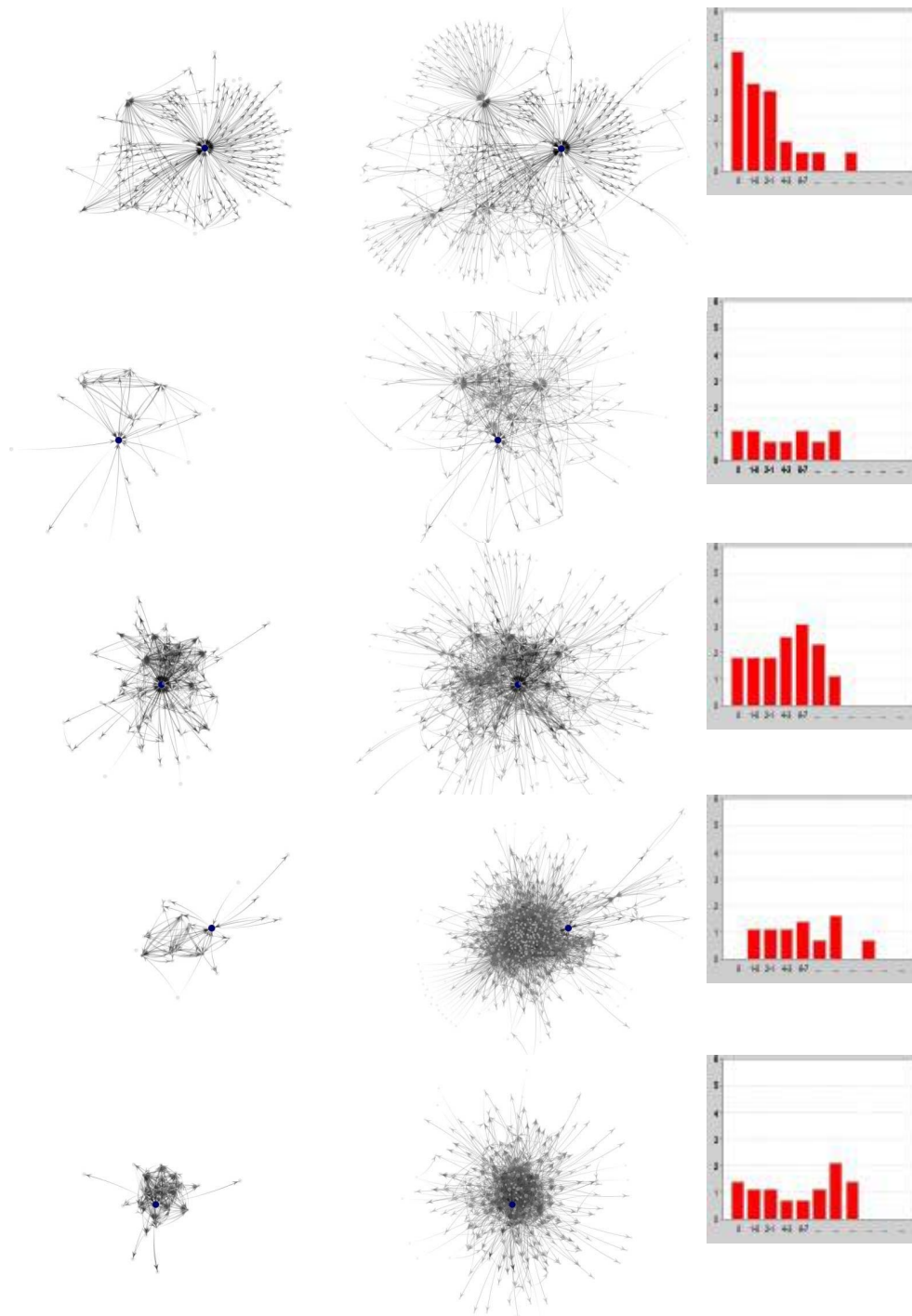


Figure 2. Reply-to network at distance 1 (left) and distance 2 (center), and out-degree distribution histogram, for one representative member for each of five newsgroups. From top to bottom, (a) *microsoft.public.windows.server.general*, (b) *rec.kites*, (c) *alt.flame*, (d) *alt.politics*, and (e) *alt.support.divorce*.

We see that this person responded largely to people of low out-degree. The graph at right is on a log-log scale, and so we see that virtually all his communication occurred with people of degree below 3, with the most at 0. This, then, is a reasonable characterization: someone who responds to many people of low degree, but infrequently to those who respond to many.

We shall see that this structural pattern is a distinguishing characteristic of question and answer newsgroups.

In figure 3, we show the average of the distributions of out-degree for the most active posters within the newsgroup—those who have been involved in ten or more replies—to emphasize that this individual is indeed fairly representative.

Contrast the *alt.flame* newsgroup (2c). This newsgroup, dedicated to the art of the flame war, involves authors flinging insults at each other on a broad variety of topics, interest areas, and degrees of obscenity and personalization. We see, immediately, that *alt.flame* looks significantly different from the technical support newsgroup.

Perhaps most visible is the density of the network. In *alt.flame*, many people connect to many others; the network is a tightly-knit one. Only slightly more subtly, we see many thick edges. In *alt.flame*, users get into intense pair-wise connections: flame wars, as the two find a topic that mutually interests them. While reading the newsgroup shows that there are often multiple participants in early stages of threads, threads often evolve (or devolve) into strong reciprocal ties. Indeed, a clear indicator of a flame newsgroup seems to be strong reciprocity. While traditionally, reciprocal turn-taking is a sign of a good conversation, we should recall that newsgroups tend to act as public spaces; the strong reciprocity is a sign of an inward focus and vigorous debate.

Note that the *alt.flame* out-degree distribution is different: it peaks at a middle value. People in *alt.flame* preferentially reply to people who are fairly well-connected: people who have responded to few people in the newsgroup are not, in turn, linked.

These groups were extremes, in opposite directions. *rec.kites* (figure 2b) is a happy medium: a discussion forum that has an in-group, but admits outsiders. Note both the well-connected core, and the less-well-connected (but largely still of greater degree, and actively participating) periphery.

alt.politics (figure 2d) is a very large newsgroup dedicated to discussion. The distance-1 graph visibly shows, again, a tightly-connected newsgroup; the distance-2 graph shows an interesting phenomenon: a population of outsiders who reply to messages within the newsgroup, but are largely not, in turn, replied to.

This suggests that the newsgroup's core has an exclusive clique aspect to it.

We can confirm this by examining newsgroup-wide statistics. Across all of Usenet, according to Netscan, 66% of posters post only one message. Ordinarily, 93% of those are initial turns: they are from people who (try to) initiate a thread. In this newsgroup, however, only 50% of those posts are initial turns: the rest are replying to messages.

While we don't have a direct explanation for how the newsgroup maintains its exclusion of outsiders, it is definitely the case that the most prolific posters clearly prefer to reply to other prolific posters. This is shown, again, by the degree distribution.

We contrast this with a very different newsgroup: *alt.support.divorce* (figure 2e), which is dedicated to social support of people who are going through or have recently been through a divorce. We have already seen that *alt.support.divorce* is different from any of the newsgroups we have discussed so far: it has a smaller percentage of authors who post only one message; it has a more equitable distribution of both in-degree and out-degree across the group. Now we will see that its most active users are also less exclusive than *alt.politics*.

Figure 2e shows a not-atypical member of *alt.support.divorce* newsgroups with the by-now-familiar degree distribution. Note the radically different degree distribution: the strong bump at the front indicates a newsgroup-wide interest in welcoming new participants, while the large bump at the back indicates a continuing preference for well-linked in-group members.

3.2. Cumulative Degree Distribution

We now begin to compare newsgroups to *each other*. In the social networks section, we examined individual out-neighbor degree distribution. We now compare these degree distributions to each other, which allows us to compare newsgroups to each other.

The histograms are shown in Figure 3. For each group, we present the *average* of the number of out-degree neighbors in each size bracket: 0, 1, 2-3, 4-7, 8-15, 16-31, 32-63, and 64-127. In order to get only authors with non-trivial histograms, we focused specifically on authors with more than ten neighbors (that is, people who either they replied to, or who replied to them): an active minority, ranging from 3% of the group's membership (in the technical groups) to 30% of the group's membership (in the flame groups), as illustrated in Table 1.

We note that these patterns reinforce what was observed above: there is a distinct preference for replying to people who had not replied to anyone for the technical support groups; there is an opposite preference in the politics groups. Support groups have both a peak at the low end and high end, representing an interest in both connecting to new people and in maintaining ties with more senior ones.

3.3. Degree Equality Within Newsgroups

Another dimension of difference that separates newsgroups from one another compares the distribution of the degrees of nodes to each other. For all these newsgroups, the distribution of vertex degree

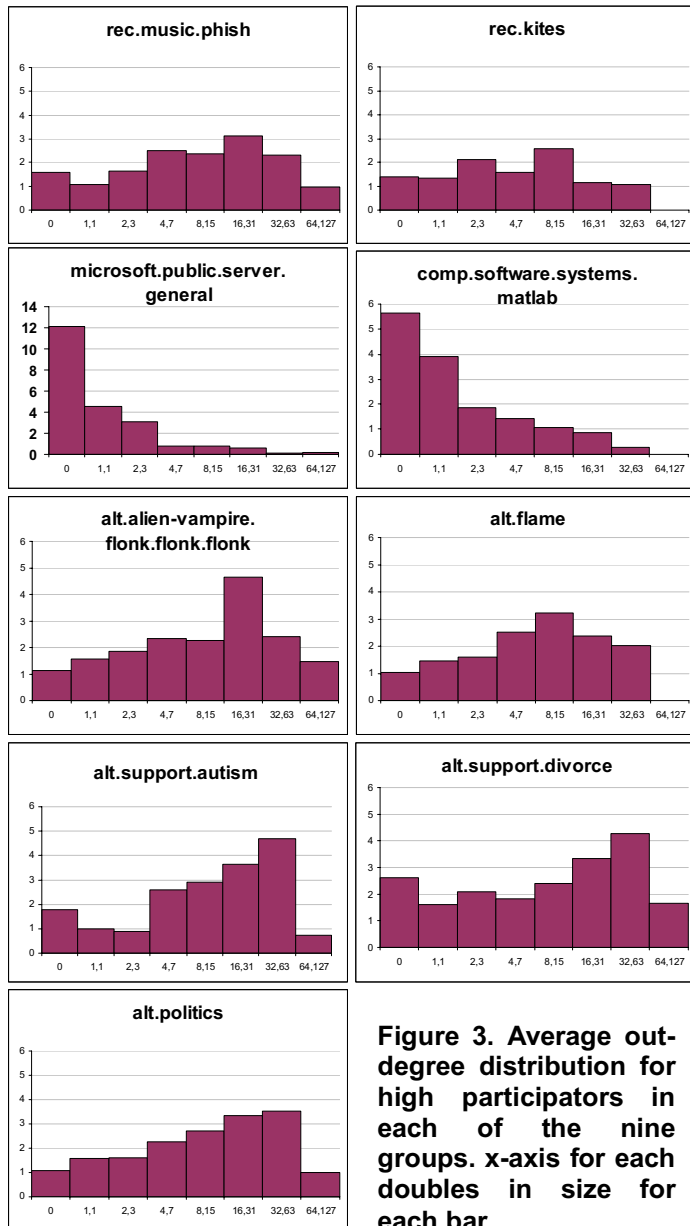


Figure 3. Average out-degree distribution for high participants in each of the nine groups. x-axis for each doubles in size for each bar.

(for degree > 0) falls along a power law: there are many at the low end, and far fewer at the high end.

The exponent on a power law can be estimated by examining the slope of the distribution, plotted on log-log axes; a negative exponent indicates a curve with lower y at higher x [1]. The nine groups are then plotted on a scatter-plot, with in-degree distribution – that is, the distribution of the number of replies that authors get – is compared to the out-degree distribution, as shown in Figure 4.

Recall that the exponent on this power law is roughly a measure of equality of contribution: a coefficient of 0 suggests a newsgroup which has a completely uniform contribution, in which each member has the same in-degree and out-degree as each other member; a strongly negative number is more inequitably distributed.

By this reading, most newsgroups fall nearly along a balanced axis, with similar distributions of in-degree and out-degree: that is, most people have the same sort of reply rate as their being-replied-to rate. The social support newsgroups, *alt.support.autism* and *alt.support.divorce*, both have a more balanced distribution, while the other newsgroups have radically different distributions.

Note also that *server.general* and *matlab* both fall substantially off of the main axis. For both newsgroups, their out-degree is distributed less equitably than their in-degree. This is, again, consistent with the model of a small population of answer people responding to a large population of questions. Virtually everyone, no matter their degree, gets a response – and so the

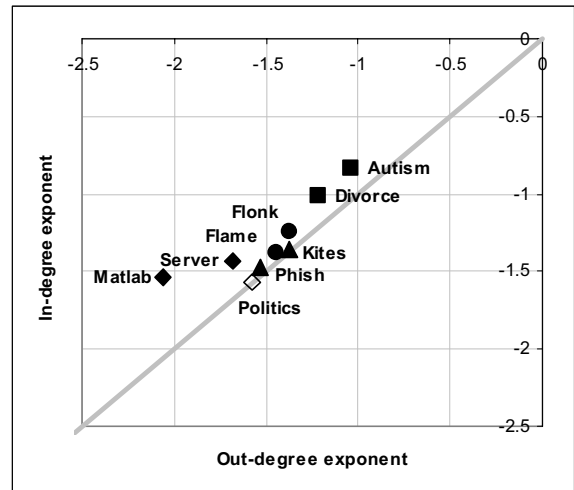


Figure 4. In-degree power-law exponent against out-degree power-law exponent for the nine groups. Shapes indicate the general group type.

curve for in-degree is fairly balanced. In contrast, the only people doing the answering are the answer people and so their out-degree is very high compared to a population with extremely low (often 0) out-degree.

It should be noted that it is hard to ascribe statistical significance for power-law measurements; thus, while these techniques are intriguing, they will require more development and precision before they can be said to be reliable.

3.4. Authors Appearing Once

One of the computationally simplest comparisons that can be made is to look at thread activity within the newsgroups. Counting the number of authors who appear only once gives a dramatic idea of how committed to the group individual authors feel.

This is illustrated most vividly in the question-and-answer newsgroups. In these, many authors will show up once to ask their question. If they get a satisfactory answer they may never be seen again. Therefore, it might be considered broadly acceptable to post a small number of times.

As further evidence that this pattern of question and answer is prevalent within specific groups, we examined a distribution of the length of threads within the groups. As table 2 illustrates, technical support newsgroups had a disproportionately high number of threads of length two, while a disproportionately small number of threads of length five or greater. In contrast, social support groups tended much longer.

4. Discussion

We have presented several techniques that allow different social spaces to be compared to each other based on patterns of social interaction within each group. There is a particular, and interesting, advantage from using Usenet data: Usenet data comes from a largely uniform set of interfaces across groups. While users may access Usenet in a number of different clients, the underlying platform is uniform. In addition, clients are not directly linked to newsgroups: thus, one would not expect everyone in a particular group to be using a specific client. This allows us an unusual degree of comparability: we can examine how one newsgroup develops against others, with no *external* cues or affordances other than the title of the group. Of course, the content of the group and the culture within it certainly *is* significant.

4.1. Roles within Newsgroups

These techniques seem to have provided some description of the behavior that occurs within Usenet newsgroups. While these have are generally accessible by “feel”—by reading the group for a long period and coming to understand it—this is a first pass at a quantitative attempt to describe how welcoming a group is to strangers, how long conversations run, and how connected members of the groups are. This, in turn, seems to partition between different types of groups nicely.

In this paper, we have examined several types of discussion-oriented newsgroups: social support, hobbyist discussion, and political discussion. In addition, we have examined flame groups and question-and-answer groups.

Discussion-oriented newsgroups have several attributes that cause them to tend toward both longer threads and toward more involvement from individuals. Discussion newsgroups are places where people attempt to establish reputations [3], get to know each other, and exchange views. Conversations, therefore, run somewhat longer. In contrast, online question-and-answer newsgroups also have a population aiming to develop a reputation, but they develop their reputation by answering questions.

As such, the newsgroup has two noticeably different populations: those who ask questions (who start conversations, send few messages, and appear infrequently in the newsgroup), and those who answer questions (who largely respond to messages that have been sent, post a great many messages, and appear continuously in the newsgroup.)

At the other end of the scale, social support newsgroups (like *alt.support.divorce*) go out of their way to greet new people. Group members also seem to want to be more participatory -- and so we see that the percentage of people who appear only once is small: 21%.

This categorization of newsgroups into “question and answer” against “discussion” is overbroad, in several ways. First, of course, conversation spaces are not monolithic: even flame groups engage in question and answer. And, second, there are many styles of discussion: conversation and social support are very different. Thus, this conversation from *alt.flame*

“Taht's [sic] like being a very juicy peace [sic] of poop”

is radically different from that in a social support newsgroup (in this case, discussing the emotional after-effects of therapy):

Yes - I had hoped that it might help, but in the end it only gave me a few pieces of paper. Which was

pretty much expected, but now I know it was right to expect it...

We have shown that these structural methods can help distinguish between types of authors involved in discussion: support, dialogue, and flame wars.

4.2. Analysis Tools for Newsgroups

We believe that the strategy of visualizing roles from local networks and structural metrics can be extended across the Usenet and to other threaded discussion spaces. For example, these techniques could be applied to online spaces that record response patterns (such as Slashdot's threads); they could also be applied to email lists. These types of structural analyses have important implications for social research, and for the ways that people use and access newsgroups.

The underlying structural principles in this analysis can be extended to the design of user interfaces. There are two important foundations for building such interfaces: first, the general concepts explored in this paper must be expanded into more precise heuristics; second, the heuristics that describe individual behavior must be aggregated to describe groups.

The pathways toward the former are clearly laid out: we have demonstrated in this paper a number of both qualitative and quantitative approaches which can seem to discriminate between different types of groups and activities. Welser et al. [14] expands on these techniques with more detailed work on identifying answer people.

Characterizing groups might be best done by aggregating data about individuals. The latter suggests an analysis of groups in which the makeup of people might be summarized: perhaps that a Q&A group is "70% questioners, 20% answers, and 10% discussion, by volume of posts." Groups could, en masse, be summarized by these sorts of statistical abbreviations.

This is desirable for the analyst, of course, but also for the user. Consider a person searching for information in a field they know little in. The techniques in this paper, if automated, could point this user toward groups that have both discussed this topic and are of the appropriate term: is the user looking for technical support, or social support? Do they want a brisk discussion, or a straightforward answer? Indeed, these techniques could lead not only toward the particular newsgroup, but to the particular persons who are likely to be of greatest interest.

This notion of characterizing groups by their emotional makeup is not a new one. This work shares it with projects like Loom [2], which—in part—attempted to color groups by their emotional content. Similarly, predecessors to this project [5, 11, 12, 14]

have done their part to characterize the content and makeup of a group. The contribution here is the start of a new set of tools, perhaps more quantitative, which can be used to break down participation in more detail.

4.3. Beyond Newsgroups

But newsgroups are not the only place where this work may be applied. While other communities may have different micro-dynamics, it seems at least reasonable to suspect that these sorts of patterns would recur within different online spaces. The back-and-forth of a flamewar, the multiparty discussion, and the hub-and-spoke of an authoritative answer person can all be used more generally. Perhaps some of these analysis and visualization techniques could be extended as "social proxies" [4] to provide indicators of the activity of different online spaces.

These patterns may also guide the community maintainer in understanding better how to interpret their online community. Depending on their goals for the community development, a transition toward—for example—ignoring new members might be not be interpreted as a pathology. Instead, it might be read as a cue that the group has formed a tight discussion core, and is moving away from a social support phase.

5. Bibliography

- [1] Adamic, L. "Zipf, Power-laws, and Pareto: a ranking tutorial." Online, retrieved on 6/14/2005 from <http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html>
- [2] Donath, J., Karahalios, K., and Viegas, F. "Visualizing Conversations." *Journal of Computer Mediated Communication*. 4(4) 1999.
- [3] Donath, J. "Identity and Deception in the Virtual Community." In Kollock, P. and Smith, M. (Eds). *Communities in Cyberspace*. London: Routledge.
- [4] Erickson, T. "Designing Visualizations of Social Activity: Six Claims" The Proceedings of CHI 2003: Extended Abstracts. New York: ACM Press, 2003.
- [5] Fisher, D. "Egocentric Networks To Understand Communication." *IEEE Internet Computing*. 9(5) 2005.
- [6] Golder, S. and Donath, J. "Social Roles in Electronic Communities." *Presented at the Association of Internet Researchers (AoIR)*. Brighton, England. Sept. 19-22, 2004
- [7] Merton, T. *Social Theory and Social Structure, Enlarged Edition*. The Free Press: New York. 1968.

- [8] Netscan, from Microsoft Research. Available at <http://netscan.research.microsoft.com>
- [9] O'Madadhain, J., Fisher, D., White, S., and Boey, Y.-B. "The JUNG (Java Universal Network/Graph) Framework." Technical Report UCI-ICS 03-17. School of Information and Computer Science, UC Irvine. Irvine, California, 2003
- [10] Sack, W. "Discourse Diagrams: Interface Design for Very Large Scale Conversations." HICSS 2000.
- [11] Turner, T., Fisher, D., Smith, M., and Welser, T. "Picturing Usenet: Mapping Computer-Mediated Collective Action." *Journal of Computer-Mediated Communication*. 10(4) 2005.
- [12] Viegas, F. and Smith, M. "Newsgroup Crowds and AuthorLines: Visualizing the Activity of Individuals in Conversational Cyberspaces." HICSS 2004.
- [13] Wasserman, S. and Faust, K. *Social Network Analysis*. Cambridge: Cambridge University Press. 1994.
- [14] Welser, H., Gleave, E., Fisher, D., and Smith, M. "Visualizing the Signatures of Social Roles in Online Discussion Groups." In preparation.