# Analysis of a Web-based Network of Educators
## Advanced Data Analysis Project

April Galyardt

Turadg Aleahmad
Stephen Fienberg, & Brian Junker
Steve Hargadon

December 16, 2009

# Contents

# Analysis of a Web-based Network of Educators

April Galyardt & Turadg Aleahmad
Stephen Fienberg, Brian Junker
Steve Hargadon

December 16, 2009

# 1   Introduction

Online tools, including social networking sites, are becoming more prevalent in teacher support and professional development. States and school districts throughout the United States are setting up online tools, such as social networks or online continuing education programs, with the idea that these tools will help teachers do their jobs better.

For example, the Illinois state board of education helped finance the Illinois New Teacher Collaborative Online to facilitate the sharing of best practices across induction programs throughout the state. Denver uses the decade-old Tapped In network to manage professional development and mentoring for groups of teacher-candidates. In South Carolina the legislature was recently asked for money to set up a Facebook-style network on Blackboard to support isolated rural teachers, although this project did not receive funding. (Sawchuk, 2008)

Ideally, dedicated online spaces would provide a place for teachers to get new ideas, reflect on instructional practices, or simply get encouragement from professional colleagues. But little is known about how teachers use these networks, and how the architecture of the site might affect interactions.

We present an analysis of Classroom 2.0, an established, active and growing community dedicated to helping teachers incorporate web 2.0 technologies into their classrooms. Web 2.0 technologies are collaborative tools, such as wikis, blogs, and podcasts that allow users to share content and/or create it together. By studying a successful network of educators, we hope to understand how teachers use online communities and what benefits they might gain from participation.

If social networks are to be a mechanism for teacher development, they must provide access to additional resources, such as expertise. This study is intended as an exploratory analysis to gain understanding of the resources a successful online professional community can provide. Guided by theory coming out of research on Communities of Practice, we focused on three questions.

First, who is interacting on the network? For example, a community that consists of teachers in their first year of teaching may be able to offer each other commiseration and emotional support, but members will not have experience to draw on when offering each other advice. We use the information provided by Classroom 2.0 members in their user profiles to create a portrait of the community.

Second, how do people use the network? Do teachers use the Classroom 2.0 community to strengthen relationships with other teachers in their own state and school district, or are they reaching out to a larger group of educators? Are they forming strong relationships with a few

individuals or looser ties between many members of the larger community? Online social networks could foster connections between diverse groups that have been shown to facilitate diffusion of ideas, and technical advice. We analyze the interactions of community members using tools from statistical network analysis.

Finally, what content is flowing through the network? It is important to understand whether educators are coming to the community for technical advice with hardware and software, or if they are coming to the community to socialize with other teachers who share similar values. These purposes along with a myriad of others, are all useful to teachers. By analyzing the content of text posted to the community, we gain an understanding of what teachers are seeking from the community. This should inform the evolution of professional development, and also the website design of future online professional communities.

Section 2 contains a review of the Community of Practice literature which guided our exploration of the interactions in the Classroom 2.0 network. Section 3 contains a full description of the data set, and the data collection process. Section 4 focuses on the members of the network, while Section 5 looks at how the members use the features of the website. In Section 6, I describe the Latent Dirichlet Allocation (LDA) model for text analysis, and use the model to analyze the content in the main areas of activity in the Classroom 2.0 network. This section also contains a full discussion of the topics discovered by the analysis.

# 2  Substantive Literature Review

## 2.1  Communities of Practice

Communities of practice are defined as groups of professionals which form to accomplish tasks and provide learning avenues. Some definitions require that "learning in communities of practice is separated from neither the activity nor the meaningful social arrangements in which the activity takes place" (Johnson, 2001). Classroom 2.0 is an atypical example of a community of practice. Members are embedded in different contexts all over the globe. The community norms of the schools that the members come from are likely to be very different. Yet they have come together for a variety of common purposes, and the community still appears to be focused on its original task of encouraging and supporting teachers who want to use web 2.0 technology in their classrooms.

One of the key concepts in communities of practice is that individual knowledge and community knowledge support each other and grow together. Specific knowledge becomes obsolete very quickly, and so communities where new knowledge can be generated, discussed and shared are a valuable asset for businesses and schools. Participation by members in communities of practice need not be equal, in fact, expert-to-apprentice relationships are common and peripheral roles can be important to community growth and learning. (Johnson, 2001)

Adler and Kwon (2002) define social capital as "the resource available to actors as a function of their location in the structure of their social relations." and note that "its effects flow from the information, influence and solidarity it makes available to the actor." In the social capital literature, four dimensions of professional relations have been identified as influential: structure of ties, trust, access to expertise, and content of interaction. (Adler and Kwon, 2002; Coburn and Russell, 2008)

## 2.2 Structure of Ties

Ties in a social network are the relationships between the members of a network. Ties are often classified as hierarchical, market based, or social (Adler and Kwon, 2002). Hierarchical ties are formed where one person has authority over another, such as between a manager and another employee. Market based ties form when products or services are exchanged for money or bartered. We assume that no member of Classroom 2.0 has authority over another member, and that no goods and services are being traded on the network, so that all network ties are social. Beyond these sociological classifications of tie structure, there are two facets of structure that we will consider: tie strength, and tie span.

Tie-strength is measured by the number of interactions in a relationship. In face-to-face networks, strong ties ease the transfer of knowledge that is tacit or not easily codified (Reagans and McEvily, 2003; Coburn and Russell, 2008). We note that strong ties in online networks may not share the same benefits if the transfer of tacit knowledge relies on gestures and other body language as well as tie strength. In contrast, weak ties play a role in diffusion of ideas, public information and technical advice (Coburn and Russell, 2008). These are exactly the sort of functions that ties in an online network could facilitate, sharing of information between loose affiliations.

Ties that span different knowledge pools are important because they facilitate access to information that may not be available in an individual's local environment (Adler and Kwon, 2002; Reagans and McEvily, 2003; Coburn and Russell, 2008).

## 2.3 Trust

Trust plays an important part in social networks, often regulating the information that flows through a network and how the information is interpreted (Coburn and Russell, 2008; Johnson, 2001; Adler and Kwon, 2002). In some papers, trust is equated with tie strength, because repeated interactions allow the formation of trust (Reagans and McEvily, 2003). We do make a distinction between trust and tie strength. Trust is what allows a community member to take risks, such as sharing an experience where they are not sure they acted appropriately, and when members act on advice, they trust that the advice given was given in good faith. The anecdotal evidence suggests that the members of Classroom 2.0 have worked very hard to create an environment where trust is nurtured, but we have no data that would allow us to quantify trust.

## 2.4 Access to Expertise

Being part of a social network gives an individual access to additional resources and information; but, the extent of additional resources available is dependent upon the expertise of the other members of the network (Adler and Kwon, 2002; Coburn and Russell, 2008). Coburn & Russell (2008) studied two school districts implementing mathematics reform, placing coaches in each school was a part of the reform in both districts. While one district promoted regular teachers with no particular experience in mathematics, the other district hired teachers who already had some expertise and then provided them with additional training opportunities. In both districts the math coaches were the source of expertise in the network, but in the district with trained coaches, all teachers had higher levels of access to expertise, even if they had no direct contact with the coaches.

## 2.5 Content of Interaction

Tie content can refer to shared norms and beliefs (Adler and Kwon, 2002), or to substance of conversation (Coburn and Russell, 2008). We restrict our interest to the substance of conversation, and focus on identifying the most common discussion topics on the network.

Coburn and Russell (2008) found that more than just information flows along social networks. "Routines of interaction crafted by the district not only flowed into schools via the coach; once in the school, they moved from teacher to teacher as well." So it is possible that information and habits brought into schools from Classroom 2.0 members then flow through the school. Discerning whether this actually occurs is well beyond the scope of this study.

## 3 Classroom 2.0 and Data Set

Classroom 2.0 is a social networking site founded to help teachers incorporate collaborative technologies, often referred to as web 2.0 tools, into their classrooms. It is an open community, and membership is free. The community platform enables multiple types of interaction between members. Each member has a homepage with a comment wall which facilitates interactions in a manner similar to MySpace. Each member also has the ability to create a personal blog, participate in discussion forums, post photos and videos, and join groups. Our data consists of a download of the content of the site from its founding in March 2007 through October 2008, including all 6 types of interactions and user profiles.

Our data includes 14,000 members, of which approximately 2000 have taken no additional actions on the network after creating their profile. One interesting phenomenon, is that around September 2008 Classroom 2.0 began experiencing a period of rapid growth. In September 2008, the network had 11,000 members, by October their ranks had grown to 13,000; in January they passed 18,000 members. Our data contains only the very beginning of this surge in membership.

### 3.1 Data Collection Process

While all of the data is public, we considered it important to have the community not only sign-off on the project, but also participate in and comment on the research. We contacted the owner of Classroom 2.0, Steve Hargadon, to both assist us with data collection and facilitate our discussion with the community. We posted an explanation of our project for the community on the site's forums[1], and held a web meeting to discuss the project with the community [2].

The web meeting was attended by about 40 people. Many of the attendees were education researchers themselves and expressed a great deal of interest in this project. They also expressed a degree of surprise that we were doing quantitative social network research instead of a qualitative study.

## 4 Who is interacting on the network?

Members of Classroom 2.0 create a user profile when they join. The fields in the profile include hometown, school/work affiliation, and an "About Me" statement. Since each member shares only

---

[1]Forum post: http://www.classroom20.com/forum/topic/show?id=649749%3ATopic%3A199474
[2]Meeting recording: https://sas.elluminate.com/p.jnlp?psid=2008-10-20.1718.M.E2778A53C1F6D563E74CF199BAC39A.vcr

the information that they are willing to make public, and that they consider relevant, there are vast amounts of missing data in the profiles. However, we are still able to glean some information about members from the profiles.

About half of the 14,000 members in our data reported their country. They come from 115 distinct countries, with the most members coming from English-speaking countries: United States, Australia, United Kingdom, Canada and India. Slightly less than half the members are Americans, and some of the most active members are Australian.

A significant majority of the network, appear to be teachers. Seventy-Two percent of profiles contain a variation of the word teacher under About Me: *teach, teacher, teaching.* The proportion of teachers may be much higher than observed because of missing data; however, context is missing. For example, we do not know whether the word *teacher* was used in the past, present or future tense: "I was a teacher," "I am a teacher," or "I'll teach when I graduate."

Among the 8000 Classroom 2.0 members that reported their gender, 57% of them indicated the member was female. According to the US Census 71% of all teachers are women. Among elementary and middle school teachers, the proportion is 79% women, while only 59% of secondary teachers are women. So there are at least 3 distinct possibilities when considering the gender proportions of the Classroom 2.0 population. It is possible that non-reporting caused a significant perturbation of the estimated proportion of women. If we assume that the members who reported their gender, do not differ from the members who withheld the information, then we can consider the other two alternatives. The site might be attracting mainly secondary teachers so that 57% female is unsurprising. Alternatively, the site may be drawing members from all grade levels, but appealing more to men so that we observe a proportion considerably smaller than the national percentage 71%.

Additionally, 11% of the profiles contain the words *college* or *university* under the Affiliation heading. However, it is unclear whether these members are professors, graduate students, undergrads, or even what discipline these members are in. Inspection of individual profiles indicates that some of those affiliated with universities are teachers who have gone to graduate school for masters degrees in teaching, others are graduate students and professors in computer science who work on education technologies. The only way to ascertain the numbers of members who truly fall within these categories would be to undertake a hand-inspection of the fourteen-thousand profiles.

The dangers in trusting public, self-reported data are illustrated in an examination of the member ages. About half of the members, 6550, reported their ages, shown in Figure 1. There are small numbers of members reporting an age of 0 or 100, but more interesting is the strange spike at the age of 35. Closer examination revealed that 200 individuals had given a birthdate of 1/1/75. One possible explanation is that this date was the default on a pull-down menu on the registration page. After these replicates are removed from the data, the distribution of ages appears to be a mixture of 2 normals centered at 34 and 52; but given that only half the members provided ages, this may not be representative of the population of members.

## 5   How do people use the network?

### 5.1   Patterns in joining the network

Many of the members joined in groups, for example; 4 people from Manhattan, Kansas joined together, later 7 from Springfield, Massachusetts, and recently 21 elementary teachers from
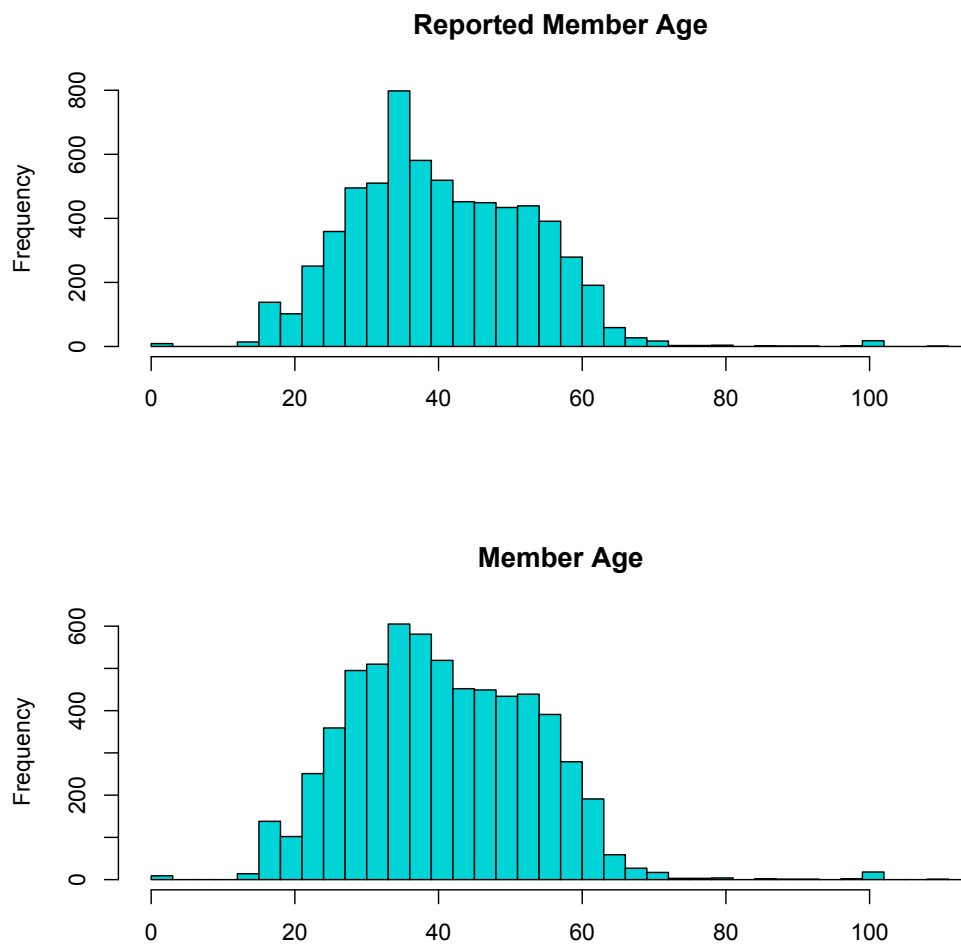
**Reported Member Age**

**Member Age**

Figure 1: The histogram on the top shows the member ages exactly as reported. The histogram on the bottom shows the ages after removing all the individuals who reported 1/1/75 as their birthday.

Wanamingo, Minnesota joined within minutes of each other. We even see one instance where a block of 50 teachers from all over Pennsylvania joined the network in a space of 10 minutes, perhaps this was a group of attendees at a training workshop.

To explore these patterns more fully, we created a visualization using Google Earth (2009). The Google Earth file is available at: http://www.stat.cmu.edu/ galyardt/Classroom20.kmz. The Google Earth software allowed us plot a point at the location of each member as they joined the network. From this we were able to observe bursts in new members joining the network. Rarely did one person from a particular location join by themselves. Instead, we see several people joining in a short time span, then perhaps several more a week or a month later. This behavior suggests that the network is growing through word of mouth, and that many members have existing relationships when they join the network.

In addition, we also observed that some cities occur more frequently as member hometowns than we might expect given the size of the cities: Manhattan, Kansas; Greensboro, North Carolina; Salina, Kansas; Colorado Springs, Colorado; Eugene, Oregon. The one thing that these places have in common is that they are home to large state universities, which have large teacher preparation programs. This corresponds with our observation that more than 10% of members are affiliated with universities and colleges. In future work, we may take a closer look at members from these locations as an interesting sub-population of the network.

## 5.2   Patterns in Use

There are 6 different ways in which members can interact within Classroom 2.0:

- Members may form colleague relationships. This relationship allows members to send private messages to each other.

- Any member may initiate a thread in the discussion forums hosted by Classroom 2.0, and they may reply to any existing topic.

- Each member has a homepage, where any member may leave comments for them.

- Each member also has a space for a blog, where only they may post, but others may comment on the posts.

- Additionally, Classroom 2.0 has photo and video galleries where members can post their own photos and videos and comment on other's photos and videos.

The number of actions taken in each area are shown in Figure 2. Activity in the Forums and Wall Comments clearly dominates the other 4 types of interaction. However, if we look at the number of members who have participated in each area, shown in Figure 3, a different picture emerges. Out of the 14,000 members, 9,000 of them have formed a colleague relationship, and many fewer members have participated in the forums or wall comment areas.

Once a colleague relationship is established, the two colleagues may send private messages to each other. Private messages are not part of our data set, so that the colleague relationships, unlike the forum and wall comment areas, carry no content in our data. Additionally, since correspondence based upon a colleague relationship is private, it is impossible to distinguish an ongoing relationship from one that has become inert. Indeed, it is even possible that the colleague relationship was
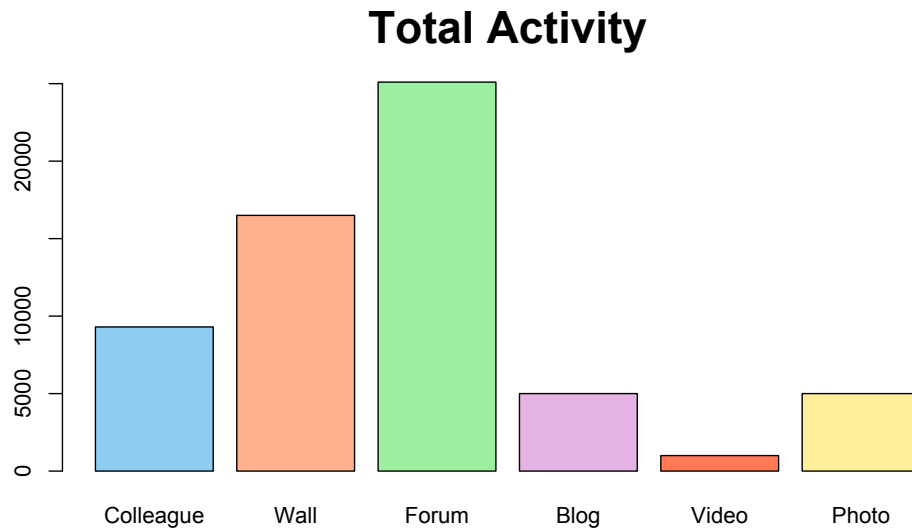
Figure 2: Histogram of the total number of interactions in each of the 6 interaction categories.
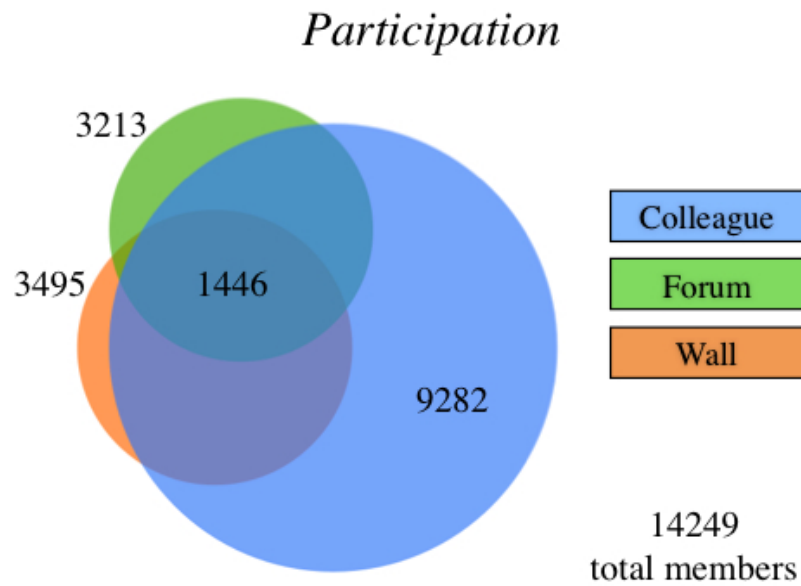


Figure 3: Venn diagram showing the overlap between people who have participated in the colleague, wall comment, and forum areas of Classroom 2.0.

9

established and that no actual interaction took place. For these reasons, we focused our analysis on the forum and wall comment interactions.

Figure 3 indicates that the network of members who participate in the forums, and the network of those who participate in the wall comments, may be largely separate networks. In the wall comment area, any comment made is left directly for a particular individual. This is intuitively a more personal interaction than the interactions in the discussion forums where an initial comment is directed to the community at large. Therefore it is of some interest whether these two areas of interaction within Classroom 2.0 foster different types of interaction.

Prior research indicates that different types of interaction have different patterns of ego networks. (Adamic et al., 2008; Fisher et al., 2006; Welser et al., 2007) An ego network is the local network surrounding one individual, the ego. It shows the ego, all of the people to whom the ego is connected, the alters, and all of the connections between the alters. Fisher et al. use and recommend a second degree ego network which looks at the network within a distance of 2 from the ego. Adamic et al. and Welser et al. both use a first degree ego network. All three papers find that areas of conversation dominated by social discussion have dense ego networks, where many approach complete graphs. On the other hand, areas of conversation characterized by expert-novice advice sharing, or question/answer forums had many stars, and almost no complete triangles. Technical experts answered the questions of many different novices, but had little contact with other experts. Novices received answers from many different experts, but had little contact with other novices.

Since the wall comments are designed to be personal and social, while the forums might serve as an advice seeking area; we theorized that we would observe dense ego networks in the wall comments consistent with socialization, and expert-novice star relationships in the forum ego networks. For ease of visualization, I constructed a random sample of 75 participants in the forums and 75 participants in the wall comments. The ego-networks associated with these participants are shown in Figures 4 and 5.

The patterns observed in the ego networks are exactly the opposite of what was predicted. The forum ego networks have many more dense, complete graphs, while the wall comment ego networks have many more stars. The content analysis in Section 6 allows us to interpret this finding. The forums are indeed being used for both discussion and question-answer type activity. However, the wall comments appear to be serving primarily as a way for members to introduce themselves to each other. After introductions are made on the wall comments, a colleague relationship could be formed in which all subsequent communication is private. Further work must be done to confirm whether contact in the wall comments tends to proceed the establishment of a colleague relationship.

## 5.3   Interactions over distance

We plotted the location of users in Google Earth (2009) and then drew the connections between users from the wall comments and the forums. The Google Earth file is available at: http://www.stat.cmu.edu/ galyardt/Classroom20.kmz An example map is shown in Figure 6. Displaying the data in this dynamic format allowed us to observe the behavior in joining the network, which was discussed in Section 5.1. It also allowed us to observe another trend: Users demonstrate a preference for interacting with network members who are geographically distant.

We observed that many members are joining the network in groups. They are joining with friends from the same school. They are joining with colleagues at workshops or conferences. When we look at the animated Google map, we can watch several people join in one town, then a few more from the same town, and then a few more. This behavior pattern, which is repeated all over

the country and all over the world, suggests that the network is growing through word of mouth.

Yet, when we look at the interactions, we see very few links between people who are geographically close. There are exceptionally few links that even occur between individuals within the same state. However, the most detailed information that we might have about a person would be city-level information, so that two users from the same city would have zero distance between them. Therefore, connections between users from the same city would not appear on the map. To determine whether this was occurring, we calculated the distance between individuals, Figure 7. There is a small mode in the data for short distances, but clearly these interactions are dwarfed by the interactions taking place over very large distances.

# 6 What content is flowing through the network?

## 6.1 Latent Dirichlet Allocation

When approaching a text analysis problem, there are two ways to think about the task. One approach would be to try to sort each document into predetermined categories; a task that could either be done by hand or automated. For example you might sort student utterances during a problem solving exercise into on-task and off-task utterances. This approach, called supervised learning, is desirable when there is some theory guiding the definition of the categories and subsequent analyses are to be performed on the sorted documents.

The second approach, called unsupervised learning, simply seeks to find patterns in the documents with no predetermined idea of the patterns that might emerge. In the case of Classroom 2.0, this was the desirable approach. We had very little idea of what the discussions might revolve around, and the few tentative ideas we ventured turned out to be badly mistaken. The other advantage of using an unsupervised learning method is that they are much less resource intensive than supervised methods. The method chosen, Latent Dirichlet Allocation, is one of the most common model-based unsupervised text analysis methods.

Latent Dirichlet Allocation (LDA) is a mixed membership model often used to discover the topics contained in a collection of documents (Blei et al., 2003). LDA uses the common "bag of words" framework in which only the presence or absence of words is modeled, and word order is ignored. LDA is an improvement over earlier models such as Latent Semantic Indexing (Deerwester et al., 1990) because it introduces a formal generative probability model. Additionally, since LDA is a mixed membership model, each document is allowed to be about multiple topics.

For example, the words *Obama, Republican*, and *Democrat* would be very likely to appear in a document from a politics topic. Similarly, words like *medicare, doctor*, and *insurance* in a document about health care. To deal with documents about health care legislation, a clustering algorithm would need to create another topic 'healthcare and politics,' but a mixed membership model is built to model exactly these sorts of combinations.

More formally, in an LDA model with $K$ topics, each topic is a multinomial distribution over the entire lexicon, with parameter $\phi_k$, so that every word appears in every topic with different probability. Each document $d$ has a membership vector $\theta_d$, where the components $\theta_{d,k}$ indicate the degree to which document $d$ belongs to topic $k$. A Dirichlet prior is placed on the parameter $\theta$. The full generative model is given with the plate diagram in Figure 8 , and the probability of a

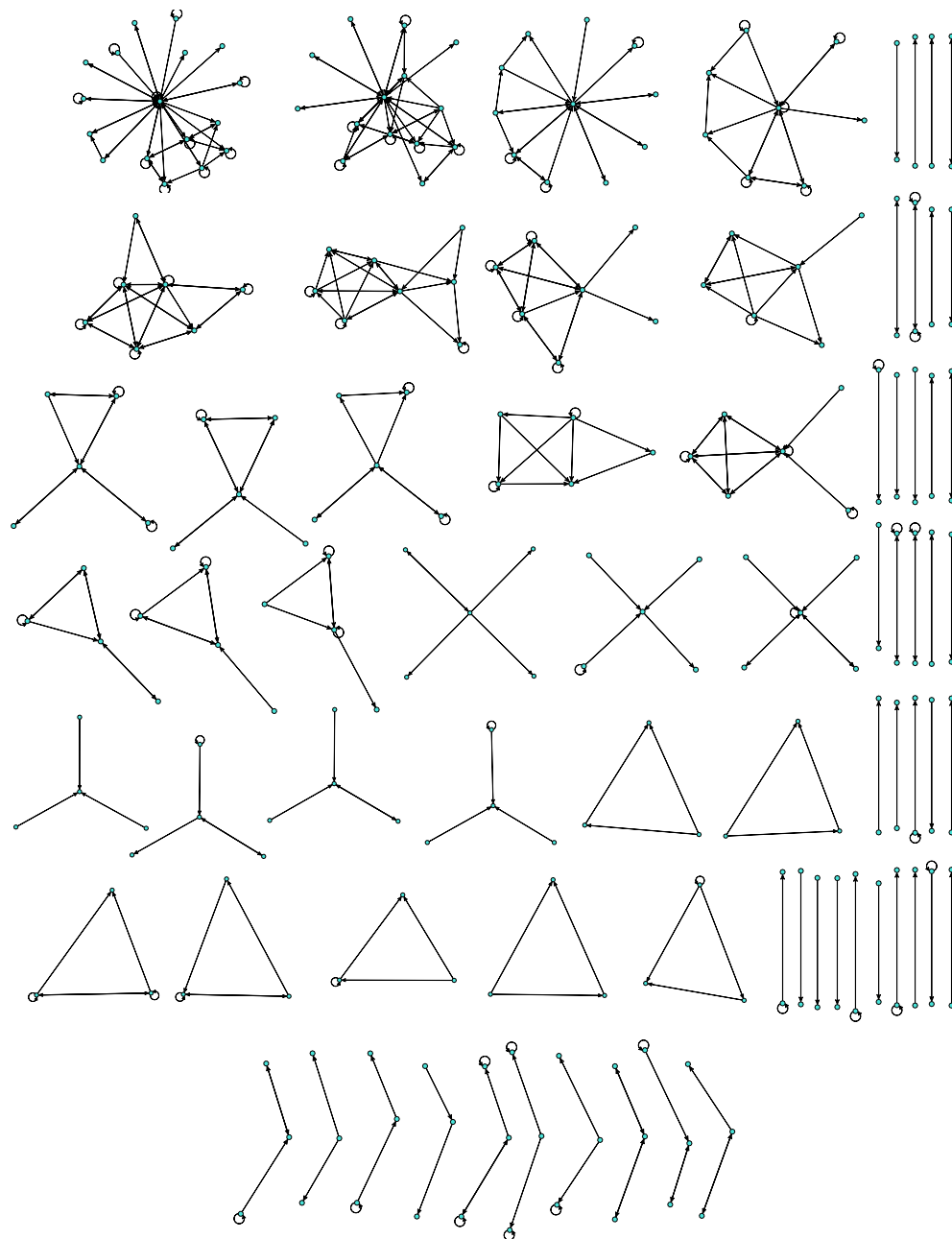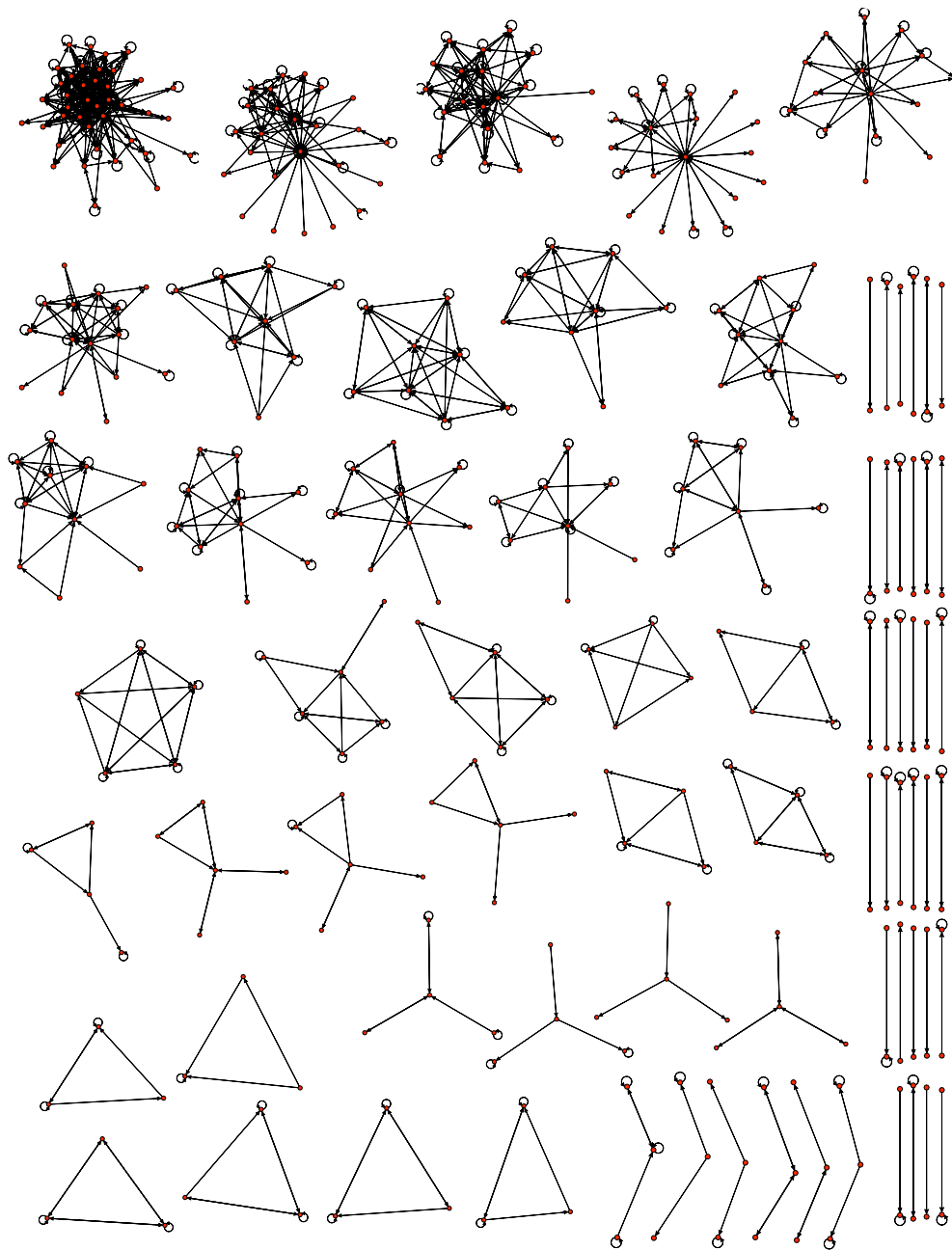Figure 4: Random sample of 75 ego networks from the wall comment area.

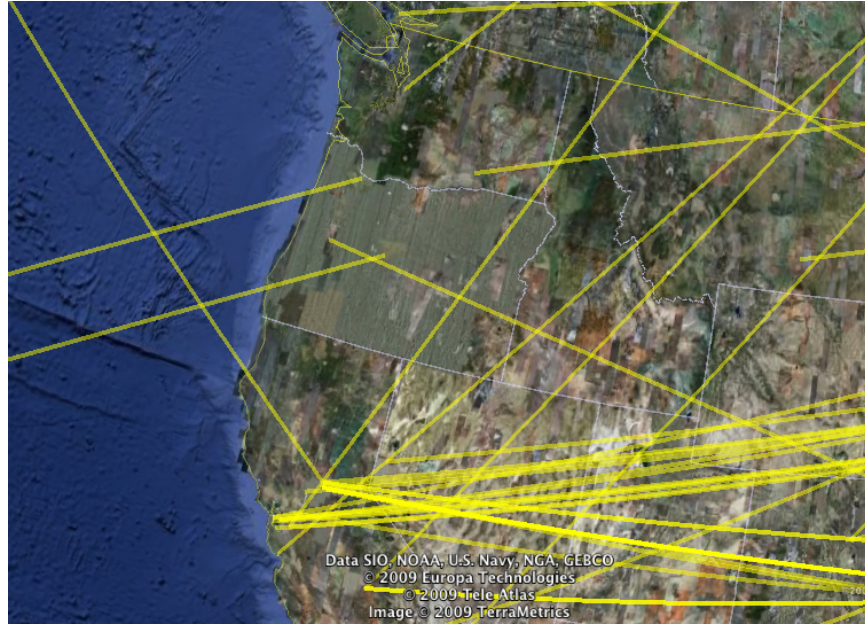Figure 5: Random sample of 75 ego networks from the discussion forums.

Figure 6: This image shows the connections between users created by commenting on each other's homepages, during a small window of time. There are extremely few local ties between users. This pattern of interaction is visible all over the world, and throughout all time windows, it is not unique to the northeast United States, nor this particular time window.
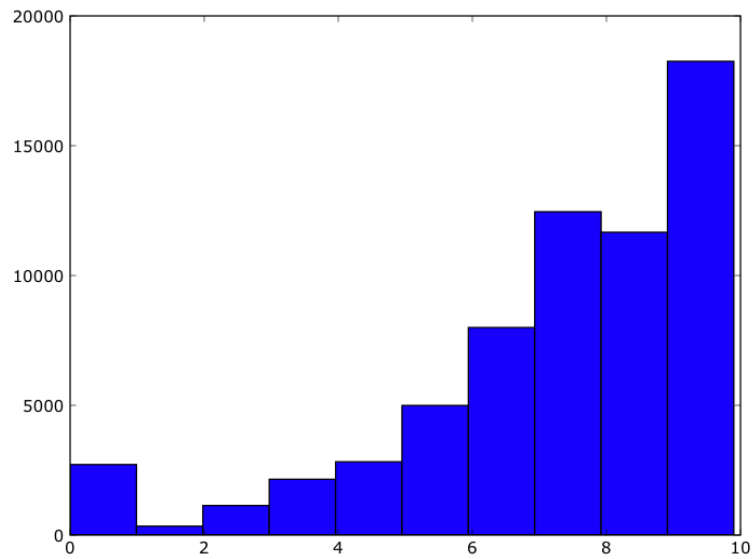


Figure 7: Histogram of the log-distance between users connected by a tie in the wall comment network. Distances are computed between cities, so that users within the same city have a zero distance. Units are log(miles).

14

document is given by:

$$p(doc|\alpha, \phi) = \int p(\theta|\alpha) \left( \prod_{n=1}^{N_d} \sum_{k=1}^{K} p(w_n|z_n = k, \phi) p(z_n = k|\theta_d) \right) d\theta$$

1. For each document $d$ in $1, \ldots, D$, draw the membership parameter $\theta_d \sim Dirichlet(\alpha)$.

2. For each of the $N_d$ words $w_n$:

   (a) Draw a topic indicator for the word
       $z_n \sim Multinomial(\theta_d)$

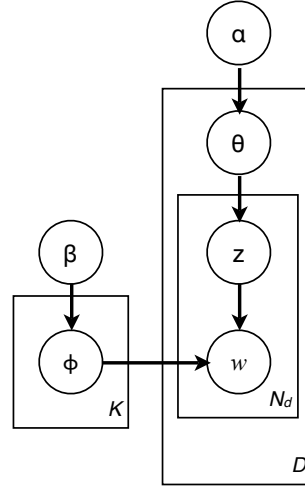   (b) Choose a word
       $w_n \sim Multinomial(\phi_{z_n})$



Figure 8: Latent Dirichlet Allocation model.

If we examine the probability of a single word within a particular document, we notice that:

$$
\begin{aligned}
p(w_n|\theta_d, \phi) &= \sum_{k=1}^{K} p(w_n|z_n = k, \phi) p(z_n = k|\theta_d) \\
&= \sum_{k=1}^{K} [\phi_{k,w_n}][\theta_{d,k}] \\
&= \theta_d^\top \phi_{w_n}
\end{aligned}
$$

Thus $w_n|\theta_d, \phi \sim Multinomial(\theta_d^\top \phi)$. This observation is essential to understanding a key feature of the model. Example 1 and Example 2 below are from the Forum data. The word 'blog' is much more likely in the *Blogging* topic, than in the *Wiki's & Google* topic, with $\phi_{Blog,blog} = 0.019$, and $\phi_{Wiki,blog} = 0.009$. So for a post like Example 1, which has high membership in the *Blogging* topic, the word 'blog' was most likely generated by the *Blogging* topic. However in Example 2, membership in *Blogging* is very low $\theta_{Blog} = 0.02$, while membership in *Wiki's & Google* is much higher $\theta_{Wiki} = 0.31$, giving:

$$
\begin{aligned}
p(w_n = \text{blog} \ \& \ z_n = Blog) &= (\theta_{Blog})(\phi_{Blog,blog}) = 0.00038 \\
p(w_n = \text{blog} \ \& \ z_n = Wiki) &= (\theta_{Wiki})(\phi_{Wiki,blog}) = 0.00279
\end{aligned}
$$

Thus for Example 2, it is more likely that the word 'blog' came from the topic *Wiki's & Google* than from the topic *Blogging*.

**Example 1. Day in a Sentence: VoiceThread-style**

*Hello everyone At my blog, I run a feature called Day in a Sentence. Teachers from around the world boil down a day or their week into a single sentence, and then post them as comments, and then I collect and publish them all on Sundays. We also have a revolving guest host system. (see Day in a Sentence posts at my blog) This week, I decided to try to use VoiceThread for the feature and I am cross-posting in a few places this week. I invite you to share a day or your week with us, using VoiceThread. You can record your voice, use a webcam or just write your sentence. (You will need an account with VoiceThread but it is worth it – it is an amazing application)*

| Topic | $\theta$ |
|---|---|
| Blogging | 0.43 |
| Announcements | 0.20 |
| Classroom Tech | 0.12 |
| Wiki's & Google | 0.01 |
| Other Topics | 0.15 |

**Example 2. Blogs, Wikis and Bloom's Taxonomy**

*When preparing a presentation to staff recently on the educational rationale for using blogs and wikis I came across this at Techlearning. Andrew Churches has done a brilliant job of linking digital technologies to Bloom's taxonomy of thinking skills. Unsurprisingly, publishing work in wikis and blogs is a great way to get students working in the 'Creating' element of the taxonomy. This is just the sort of work we need to bring out the why of wikis and blogs rather than just the how. Check it out!*

| Topic | $\theta$ |
|---|---|
| Wiki's & Google | 0.31 |
| Classroom Tech | 0.11 |
| Teachers, Tech, & PD | 0.10 |
| Blogging | 0.02 |
| Other Topics | 0.46 |

### 6.1.1 Technical Details

To fit the LDA model, I used the Gibbs sampling implementation by Phan and Nguyen (2008). No modifications to the Gibbs sampler were required; however, the output labeled "final model" by the code was simply the sample from the final iteration. I was able to calculate the posterior mean for $\phi$ and $\theta$ by saving the results from every tenth iteration, in effect thinning the chain. The hyperparameters $\alpha$ and $\beta$ were fixed to $\alpha = 0.5$ and $\beta = 0.1$.

When processing text data for analysis, it is common practice to remove 'stop words' from the data, since removing the most commonly used words in language can improve processing speed. It is generally believed that removing these words does not affect the analysis since these words appear so frequently that they would be high probability words in all the topics. Our data set was small enough that it was not necessary to remove the stop words for the sake of processing speed. Moreover, it appears that in this data set, including the stop words might have affected the topics discovered.

For example, one of the topics discovered in the wall comments we dubbed *Hi, nice to meet you.* This topic placed high probability on the words *you, I, to, hi, see, a, great, your, it, hey, in, here, good, are, hope, nice, what, meet, just, glad,* and *love.* More than half of these high probability words are stop words, the other half are politeness modifiers: *great, good, hope, nice, glad, love.* At a bare minimum, removing the stop words would drastically change the interpretation of this topic. Another possibility is that after removing the stop words, the remaining words do not hang together well enough for the topic to be modeled.

For this reason, I did not remove any words from the data set prior to the analysis. It may be found in future analyses that removing stop words does not change the analysis of text from formal sources, such as news articles, while the analysis of text from informal communication, like Classroom 2.0, is more affected.

### 6.1.2   Model Selection

Model selection is critical to obtaining an interpretable set of topics. Choosing a model with too many topics, can result in topics that pick out superfluous patterns in language. For example, when we tested a model with 50 topics on the Forum data, one topic identified first-person pronouns, another identified third-person pronouns, and yet another topic identified past-tense time indicators, such as *last, week, back, when, before.* These are real patterns in language, but clearly they do not lend to interpreting the topics of discussion.

In many settings, the preferred method of model selection would be a likelihood based method, such as Akaike information criterion (AIC) or Bayesian information criterion (BIC). However, in LDA the likelihood is computationally infeasible:

$$p(doc|\alpha, \phi) = \int p(\theta|\alpha) \left( \prod_{n=1}^{N_d} \sum_{k=1}^{K} p(w_n|z_n = k, \phi) p(z_n = k|\theta_d) \right) d\theta$$

One method that is often used to deal with this integral is approximate variational inference (Bishop et al., 2003; Braun and McAuliffe, 2007), but since we used Gibbs sampling to fit the model, using variational inference for model selection was less than ideal. Variational inference uses a simplified version of the likelihood to estimate the model parameters, often yielding biased estimates. Gibbs sampling is slower, but it does result in better estimates of the parameters and their distributions. Instead I devised a new method, whose theoretical properties should be established in future work.

Gibbs sampling provides a set of draws from the posterior distribution of the parameters. Thus if it is feasible to calculate the probability of a data point given the parameter estimates, it is also feasible to obtain an estimate of the average likelihood under the posterior distribution.

$$\mathbb{E}_{\hat{\theta}, \hat{\phi}} \left[ p(corpus|\hat{\theta}, \hat{\phi}, \alpha) \right] = \mathbb{E}_{\hat{\theta}, \hat{\phi}} \left[ \prod p(doc|\hat{\theta}_d, \hat{\phi}, \alpha) \right]$$

As noted previously, the probability of a document given $\hat{\theta}$ and $\hat{\phi}$ is simple to calculate:

$$p(doc|\hat{\theta}, \hat{\phi}, \alpha) = \prod_{n=1}^{N_d} \sum_{z_n} p(z_n|\hat{\theta}) p(w_n|z_n, \hat{\phi}) \quad which \ is \quad Multinomial(N_d, \hat{\theta}^\top \hat{\phi})$$

Since $\phi$ is a vector with length around $30,000$, the individual terms of $\phi$ are incredibly small. This could lead to numerical instability calculating $p(doc|\hat{\theta}, \hat{\phi}, \alpha)$, and even more instability when

multiplying those terms to calculate $p(corpus|\hat{\theta}, \hat{\phi}, \alpha)$. Thus it was necessary to work in log-scale, and calculate:

$$\mathbb{E}_{\hat{\theta}, \hat{\phi}} \left[ \log p(corpus|\hat{\theta}, \hat{\phi}, \alpha) \right] = \mathbb{E}_{\hat{\theta}, \hat{\phi}} \left[ \sum \log p(doc|\hat{\theta}_D, \hat{\phi}, \alpha) \right]$$

To verify that this model selection criteria chose an appropriate model for the forum data, I corroborated the choice by inspecting the model results to determine which topics were stable for different choices of $K$. As shown in Figure 9, the model selection criterion does not change substantially for $K > 20$, but there is a dramatic increase between $K = 15$ and $K = 20$. Thus the model selection criterion indicates that $K = 20$ is an optimal model for the Forum Posts.
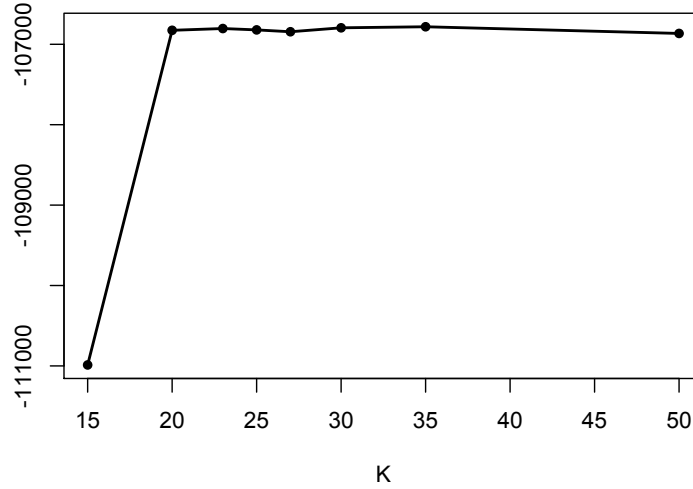


Figure 9: Model selection criteria, $\left( \mathbb{E}_{\hat{\theta}, \hat{\phi}} \left[ \log p(corpus|\hat{\theta}, \hat{\phi}, \alpha) \right] \right)$, for the original posts in the Forum Threads.

Inspection of the topics identified at different levels of $K$ also indicates that 20 is an optimal number of topics. It is common in latent class models that when a model is over-fit classes will be duplicated or split. This pattern is exactly what I observed for $K > 20$, and two examples are given in Table 1. However, for the models with $K < 20$, all stability in the topics disappeared. There were no common topics identified for the models with $K = 15$ and $K = 20$. This observation agrees with the model selection criterion that $K = 20$ is the appropriate number of topics to fit to the forum data.

## 6.2   Forum Content

Posts to the Classroom 2.0 discussion forums are often long and in-depth. The average initial post to the forums is the length of an abstract, about 125 words. Threads in the forums average 5 or 6 replies, though some threads dedicated to sharing blog links, or introductions are much longer with hundreds of replies (Figure 10).

LDA analysis of posts to the forums indicates that teachers are seeking advice in a variety of areas, both related to the Web 2.0 technology to which the community is dedicated and related to
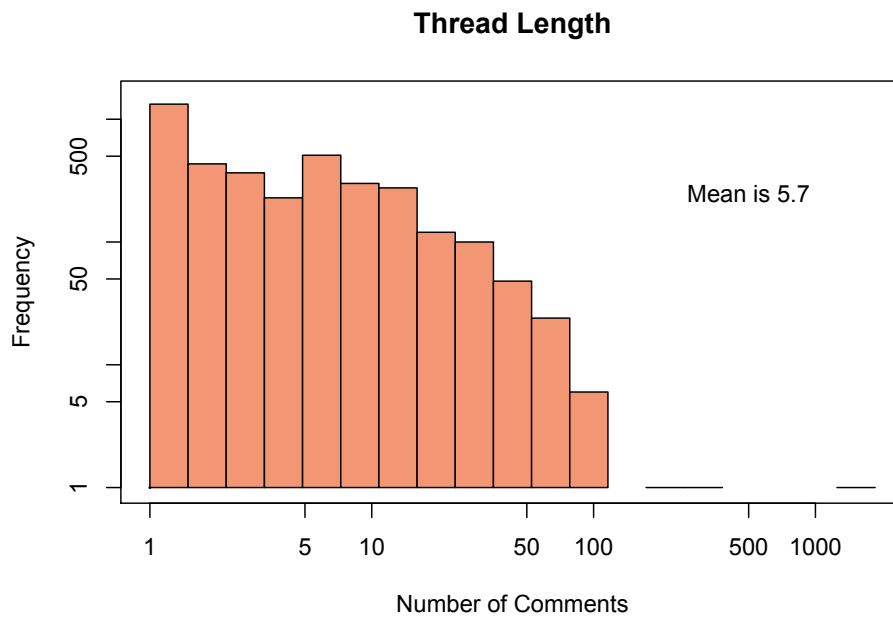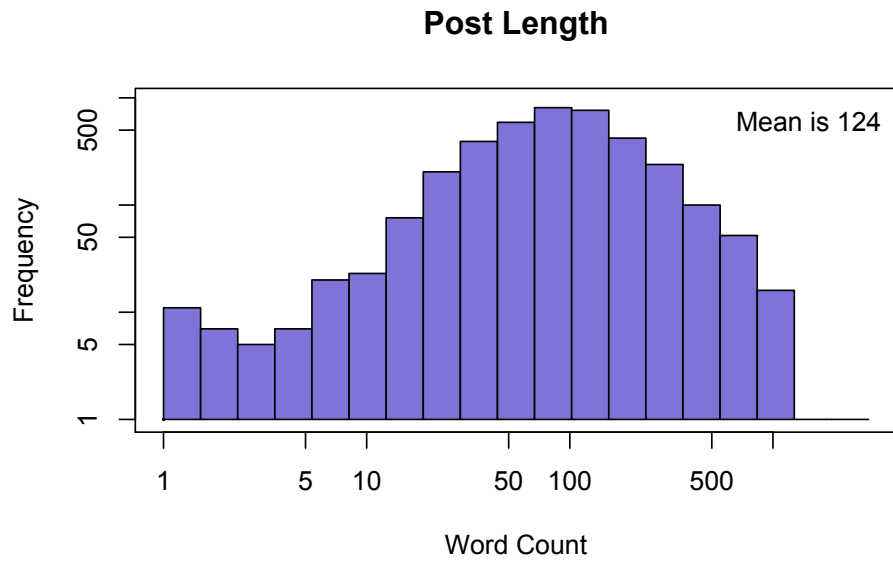
Figure 10: Histogram on the top shows word count of initial posts to the discussion forums. The histogram on the bottom shows the number of replies each post received. In both histograms, the $x$ and $y$ axes are shown on log-scale.

Table 1: The words listed are a few of the high probability words identified for each topic. On the left is a topic which was duplicated at $K = 30$, but was identified as the same topic at lower $K$. On the right, is a topic which was split into two topics at $K = 25$ and $K = 30$, but was again identified as the same topic for $K = 20$.

| K =20 | K=25 | K=30 |
|---|---|---|
| video | video | video |
| powerpoint | powerpoint | podcast |
| presentation | presentation | music |
| file | podcast | audio |
| audio | audio | youtube |
| youtube | youtube | powerpoint |
| flash | file | video |
| dvd | convert | presentation |
| convert | format | file |
| download | download | convert |

| K =20 | K=25 | K=30 |
|---|---|---|
| have | software | software |
| anyone | open | open |
| software | source | using |
| school | computer | computer |
| open | apple | source |
| computer | laptop | laptop |
| source | technology | school |
| moodle | school | district |
| experience | teachers | technology |
| district | district | teachers |
| laptop | student | student |
| computers | use | access |

more low-tech concerns such as books for story-time and classroom games. Members are also using Classroom 2.0 to find partner classes for projects such as podcasting. Some of the more interesting topics discovered are discussed below, while all 20 topics are listed with some of the informative high-probability words are in Table A in Appendix A.

### 6.2.1   Classroom Help Seeking

The topic we titled 'Classroom Help Seeking' was characterized by a list of words that are almost automatically assembled into phrases as you read the list. Phrases such as: *I am looking for some ideas*, and *I would like to... with my students*, or *I would love some suggestions.* Reviewing posts with high membership in this topic verified that this intuition was correct. Example 3 is typical of posts with high membership in this topic.

Teachers are looking for new ideas, or advice in their classroom practice. As in Example 3, they may be looking for new ideas that do not involve technology. Alternatively, a post that is high in both the Classroom Help Seeking topic and the Wiki's and Google topic, might be looking for suggestions on using wiki's to teach a particular lesson. This topic is important because it provides evidence that teachers are using Classroom 2.0 to directly affect their classroom practice.

**Example 3. Density Labs in Middle School**
58% membership in Classroom Help Seeking

> *I am new to classroom 2.0 and I was hoping to get a little help. I am tired of doing the same old density lab with my 6th grade science students. Does anyone have any ideas for a density lab that would work for 6th graders? Your help would be great.*
>
> *Thank you*
> *\*Signature\**

### 6.2.2 Technical Help Seeking

It should not be surprising that in a community dedicated to helping teachers incorporate web 2.0 technologies in their classrooms, we find a content topic dedicated to asking for technical help. This topic places high probability on words such as *software, school, open, source, computer, use*; but it also places high probability on words like *has, anyone, used, what, been, experience.*

Educators are not just asking, 'how do I do this?' Rather they are looking for suggestions on which software to choose, how to get funding for computers, and wondering which open source or freeware programs work well. The posts with high membership in this topic illustrate how teachers and schools are trying to work around budget constraints, and they are coming to Classroom 2.0 to do product research.

**Example 4. Open Source Office Suites**
40% membership in Technical Help Seeking

> *My district is exploring Open Source solutions as a cost saving measure. We are using Microsoft Office at this time. I have looked at Open Office but don't know where else to go. I need a solution that is supported on some level. We are using Drupal to redesign our district website and would like to explore more solutions for our productivity software. I have looked at Google Docs and Zoho also but I am told they don't want a web-based solution at this time.*

> *We are also looking at Linux as a possible OS solution. The version we are looking at so far is Ubuntu. We have to have a solution that works with Windows products, we use Lexia and Read 180 reading software and have a variety of other products that we have to use that are Windows compatible.*

> *What other solutions or ideas can you give us, is anyone out there using Linux in their district now? What are some good versions for a district such as ours?*

### 6.2.3 Specific Tools

Several topics were dedicated toward specific web 2.0 technologies: 'Blogging,' 'Wiki's & Google,' and 'Presentation Media.' Blogging appeared affiliated with a wide variety of subjects. Very often posts with high membership in Blogging were made by a teacher who had been using blogs as a student assignment, and the teacher was now inviting other teachers to comment on the students' blogs. This pattern is observed in Example 5. I also observed that blogging seemed to be heavily favored by writing teachers, as a way for students to practice.

All by itself, Google's prominence in Classroom 2.0 topics is very interesting, though in most cases, the subject of discussion was not the google search engine. Rather teachers were considering how to use the myriad other google tools in their lessons. Is Google Documents a good tool for working on word processing? How easy is it to incorporate Google Earth into a social studies class? Even, sharing success in using Google Scholar as a research tool with high school students.

In this context, it seems reasonable that Google and Wiki's combined into a single LDA topic. Wiki's were also being used for many of the same purposes. That is, creating alternative assignments and assessments, compared to traditional paper and pencil. For example, a common type of post would have the class as a whole create a wiki page about their current topic, anything from ancient Rome, to pond ecology.

Finally, there is the Presentation Media topic. This topic placed high probability on the words: *video, powerpoint, presentation, audio, youtube, flash, dvd, ipod, convert, podcast.* Many posts with high membership in this topic focused on technical questions such as converting one type of media to another, or embedding video in a powerpoint presentation. Many other posts featured teachers bragging about a video their class had posted to youtube, and asking other teachers to check it out. Additionally, I found many posts where a teacher was seeking a partner classroom to collaborate on podcasting. These collaborative podcasts seemed to be particularly favored by foreign language teachers.

### Example 5. School Life Collaboration Project
42% membership in Blogging

> *I've finished making a project for students to share how they feel about different aspects of their school life. The project starts with a series of blog posts that students are invited to contribute moderated comments to. They can then make a presentation about how they feel about their own school life. It would be great if some students from different countries could add some comments to the blog posts so that they can see how school life differs in different countries. If you would like to take part visit www.mytree.notlong.com*
>
> *If you have any questions and/or feedback, feel free to email me via ning.*
> *\*Signature\**

### Example 6. Internet scavenger hunt
43% membership in Wiki's & Google

> *Greetings all. I am working on a teaching unit about effective internet searching strategies. I wanted to create a fun culminating activity in the form of an Internet scavenger hunt. The students will have one class period to try to find the answers to a long list of questions. I would like them to use a variety of sources instead of 100% Google. Can anyone direct me to a good pre-existing Internet scavenger hunt that I could adapt to my lesson?*

### Example 7. Flip
42% membership in Presentation Media

> *I want to use a Flip video camera to take videos of my track team for teaching purposes. We take the videos at practice and throw them in a shared drive at school. Then kids can view during study or during a video review practice.*
>
> *Flip requires the codec on the machine or in the host computer in order to view the movies. Anyone know how to view the videos on a shared drive as we can with any other camera?*

### 6.2.4   Collaboration

The Collaboration topic largely focused around seeking collaboration partners for different projects. This topic in particular illustrates the benefits of mixed-membership modeling. Teachers are looking for collaboration partners in all sorts of projects. So that the Collaboration topic quite often appeared with the Presentation Media, Blogging, and Foreign Language topics, among others.

In addition, this topic uncovered a need within the community. Members are posting requests for collaboration partners on the Forums, but this sort of search could perhaps be better facilitated by a different feature in the website. Steve Hargadon, the founder of the community is planning development of just such a feature.

**Example 8. Simple collaboration?**

52% membership in Collaboration

> *We are looking for students from across the world to collaborate on Projects and enquiries centred around social studies and citizenship. Our students are aged 11-16 and in the next few months we move into a building which makes web2.0 collaborative work a real possibility. We are keen to work with schools who have an interest in exploring the diversity of cultures between collaborative schools and then embark on exploring key scoial issues from our own cultural perspectives. I'm rather new to all of this but am passionate about providing our students with a voice to be heard outside the bubble of our local community and anticipate some powerful learning experiences if involvement from a far off school becomes a reality. Please get in touch if you think that this might interest you or your school.*

## 6.3 Wall Comments

The LDA model indicated the presence of 15 topics within the Wall comments. Of the six interpretable topics, five were personal greetings. These topics were titled: "Welcome," "Hi, nice to meet you," "Thanks for your comments," "Check out my link," and "Let me tell you about myself." (See Examples 9-13.)

The implication is that while the Forums are being used for substantive communication centered around teaching, the Walls are being utilized to initiate relationships with potential professional friends. We believe that such contact may initiate the formation of a colleague relationship, so that all future communication is private and invisible.

Five of the topics found in the Wall comments were identifiable as spam. Though, perhaps oddly, all of the spam observed was on-topic, and perhaps even useful spam. This phenomenon is nicely illustrated by the the topic which we dubbed "Greetings from Germany." One user, Hans, posted 1086 wall comments in our data. The content of his numerous posts was so uniform, that the LDA model fit him with his topic. We hand inspected more than thirty of is posts, each of which was almost identical to Example 14. They are personally addressed to the recipient, with "Greetings from Germany" offered, Hans then issues an invitation to join his group Digiskills. The posts were clearly written one by one, despite the overall spam-like effect. Hans' ego network is shown in Figure 11, and though it is an extreme example, still fits the pattern of starred ego networks within the wall comments.

**Example 9.**

65% membership in "Welcome"

> *Hello \*name\*,*
> *Welcome to Classroom 2.0. This is a great place to share and learn.*
> *\*Signature\**

**Example 10.**

53% membership in "Hi, nice to meet you"

*It was nice to meet you today (and to see your Mini – here's ours:*
*\*Picture of a Mini\**
*I hope you enjoyed the workshop. Stay in touch.*

**Example 11.**

51% membership in "Thanks for your comments"

*Hey Kevin - thanks for the comment! I found out about your Driving Questions podcast by reading something on Ginger's site. I've watched to your last three episodes and was very impressed and motivated by a few things you said (especially the bit where you said that we've ALL got something to contribute). So, thanks for that.*

**Example 12.**

52% membership in "Check out my link"

*\*name\*,*
*My blog is at www.talentedandgifted.net. My students blog at www.giftedkidsnetwork.com/wp*
*\*Signature\**

**Example 13.**

64% membership in "Hi, let me tell you about myself"

*Hello! I am a 3rd grade teacher from Illinois and I'm trying to connect with other teachers and academic technology folks.*

**Example 14. Greetings from Germany**

*Hi \*name\*, my greetings from Germany! I focus on Web 2.0 apps in the classroom. Feel free to check some of my sites. And I've created the biggest group on CR 2.0 called DigiSkills with now 361 members. Would love if you decide to join this group about digital teaching methods. Hans*

## 6.4  LDA and Spam

Substantive discussion topics discovered by the LDA model had remarkably different posterior distributions of the membership parameter $\theta_k$ than topics dominated by spam. For substantive discussion topics, the distribution of $\theta_k$ has a single mode near zero and a long right tail. Spam topics have a bimodal distribution, where $\theta_k$ is either very small or very large. This is shown in Figures 12 and 13.

I believe that this distribution of membership in spam topics is driven by the nature of spam itself. Identical, or nearly identical documents are repeated a large number of times in the corpus. These documents may contains sets of words that do not usually appear together, such as *activity, toolkit, smart, technologies, create, customized*, and *smartboard*. Yet, the parameter estimates for the model must attempt to account for a set of documents where these words always appear with each other. So one of the extreme profiles in the model is required to account for the spam, but since no other documents share this pattern of words, the membership parameters, $\theta$ for this extreme profile, will either be near 0 or near 1.

While this may be a useful method for identifying some spam with unusual patterns of words, other spam with less unusual combinations will blend into the background. For example, "Hi, I'm from China, and I'm looking for someone to help practice my English." could appear to be a mixture of the Foreign Language and Help Seeking topics.
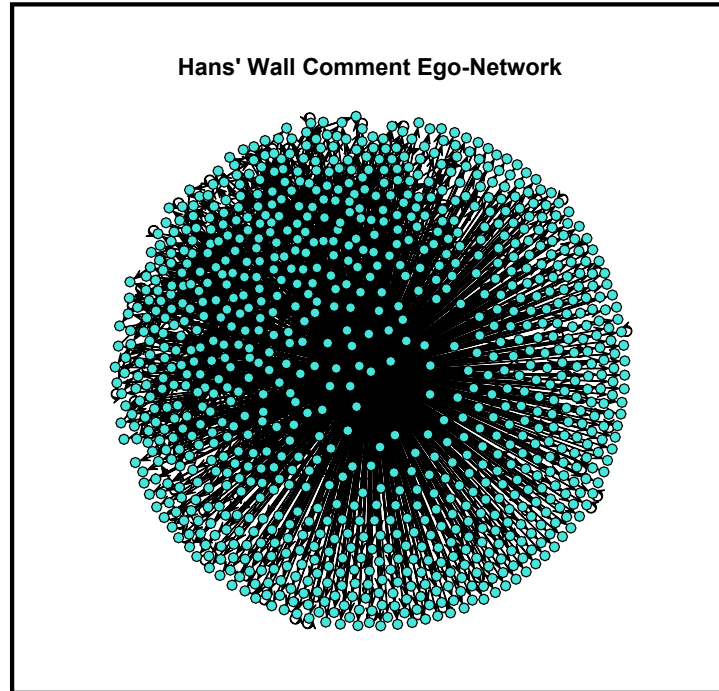
Figure 11: Ego network for Hans, the author of the "Greetings from Germany" wall topic.

# 7  Discussion

We used Adler and Kwon's (2002) definition of social capital, "the resource available to actors as a function of their location in the structure of their social relations." Our analysis has highlighted many of the resources that are available to the Classroom 2.0 members as a function of their participation in the community.

Classroom 2.0 provides relevant content for teaching using technology. Technical help is available for a wide variety of software and hardware. Help integrating the technology into a classroom environment is also available. In addition there are a wide variety of discussions revolving around whether and how technology will change education. We emphasize that the content available is *relevant*, whether such content is *useful* for informing classroom practice must be established in future work.

The different features of the site are promoting very different types of behavior. The Forum discussion boards are set up so that replies are made to a topic, while in the Wall comments replies are made to a person. Members who joined for information and discussion about education may be drawn to the Forums, while members who joined for professional camaraderie may find the Wall network more inviting. The two networks appear to have a small amount of overlap in participants, and the content differs greatly.

Substantive discussions are much more prominent in the Forum area of Classroom 2.0, while the Wall comments featured personal introductions and a wide variety of 'on-topic' spam. The forums were also encouraging denser networks, where discussion flows between all the members of the network. This contrasts directly wall comments which were dominated by star-shaped ego-networks. In the wall comments, conversations in which 3 or more people participated were relatively absent.
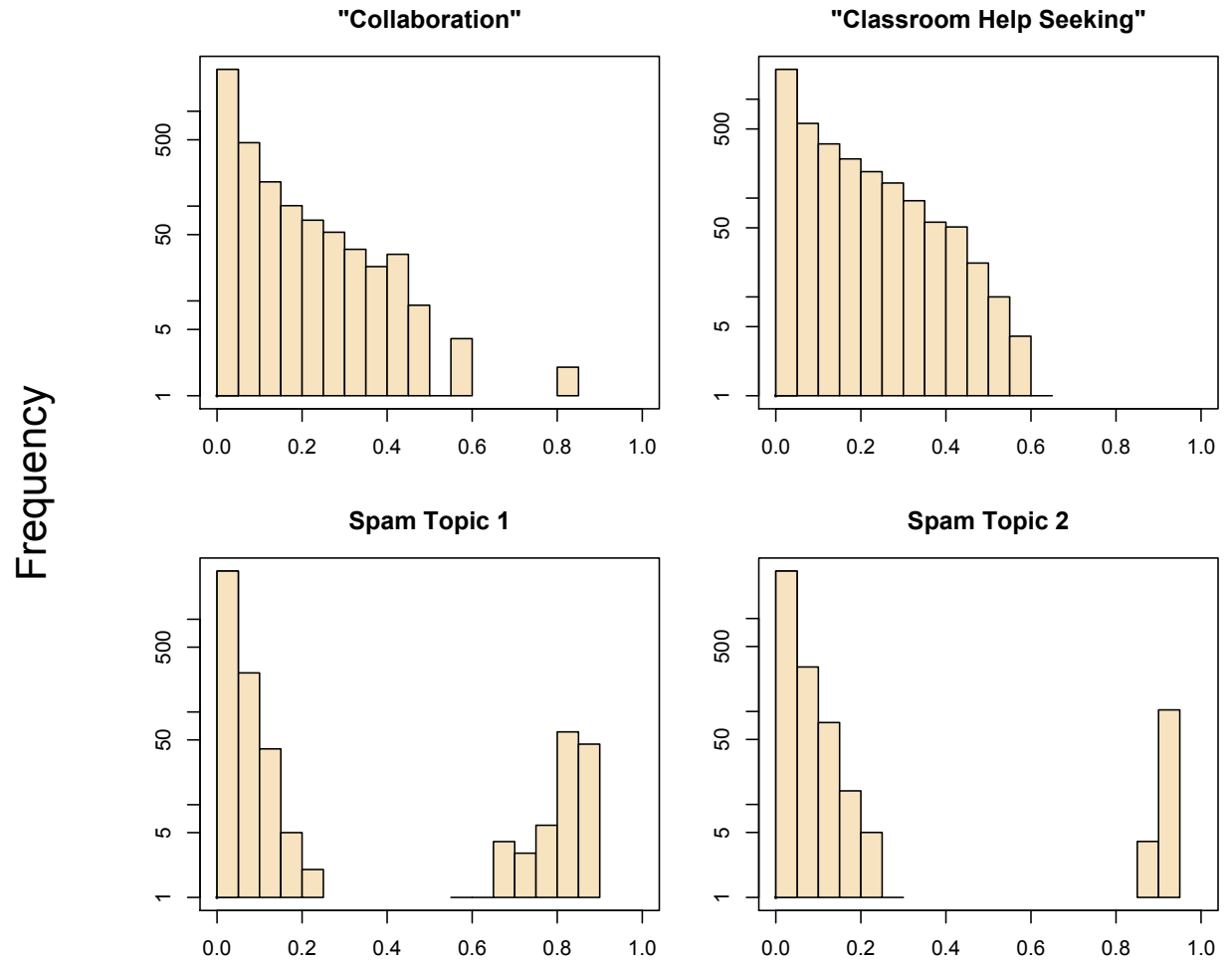
Figure 12: Histograms of the membership parameter, $\theta_k$, for posts in two substantive Forum topics and in two Forum spam topics. Frequency is shown in a log scale.
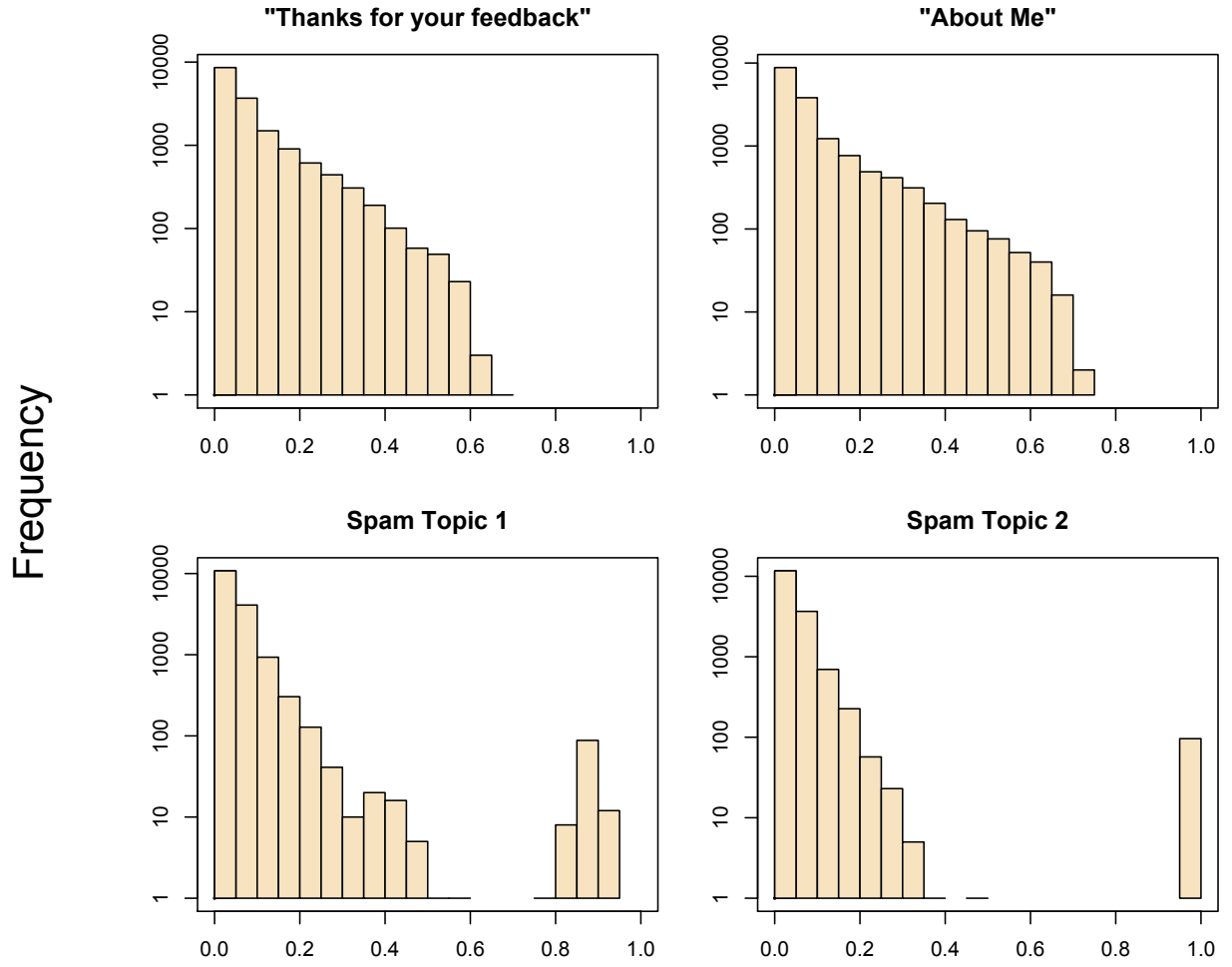
Figure 13: Histograms of the membership parameter, $\theta_k$, for posts in two substantive Wall Comment topics and in two Wall Comment spam topics. Frequency is shown in a log scale.

We believe that the Wall comments were used primarily for introductions which then could lead to the formation of a colleague relationship. Future work is planned to explore whether this theorized relationship between communication in the Wall comments and establishment of a colleague relationship holds.

We observed teachers using Classroom 2.0 to share work their classes have done, for example through sharing links to videos and wikis. They are also using Classroom 2.0 to seek out partners for future classroom projects. Indeed, the discussion forms were a very popular place to seek out collaboration partners, but the structure of a forum may be less than ideal for finding a partner classroom. For this reason, in the next generation of the community, Steve Hargadon intends to develop a better online setting for setting up collaborations between teachers.

It is worth noting that Foreign Language and Writing are highly visible in the community, while other subjects such as History, Social Studies and Art can also be observed. However, Math was effectively absent. It is unclear whether Math teachers are not present in Classroom 2.0 because they are drawn to other communities or because they are not incorporating collaborative technologies into their classrooms.

Previous work has shown that weak ties in social networks and ties that span different knowledge pools facilitate access to information not easily available locally and play a role in diffusion of ideas, public information and technical advice, (Coburn and Russell, 2008; Adler and Kwon, 2002; Reagans and McEvily, 2003). This study indicates that Classroom 2.0 may be facilitating this process. The majority of connections in Classroom 2.0 are over vast distances, connecting members in geographically diverse locations. A significant proportion of the Classroom 2.0 community is affiliated with colleges and universities, possibly increasing the diversity of the knowledge pools available. Moreover, the content analysis indicates that much of the discussion does center around public information and technical advice. We note that since members seem to prefer interacting with educators outside of their local network, schools and districts may want to consider carefully the implications of creating an online tool intended solely for local teachers.

Finally, the extent and quality of the resources available to an individual through their social network depends upon the expertise of the other members of the network. There remains some question about what expertise is available through Classroom 2.0. We know that University faculty and IT professionals are members of the network, so there is an incredible potential for sharing expertise with teachers. Due to the vast amounts of missing data in the member profiles, it is difficult ascertain the overall activity for these members. The most important component of future work is to explore the expertise of the most central members of the community.

# References

(2009). "Google Earth 5.0."
  URL http://earth.google.com/

Adamic, L. A., Zhang, J., Bakshy, E., and Ackerman, M. S. (2008). "Knowledge sharing and yahoo answers: everyone knows something." In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, 665–674. New York, NY, USA: ACM.

Adler, P. S. and Kwon, S.-W. (2002). "Social Capital: Prospects for a new concept." *Academy of Management Journal*, 27(1): 17–40.

Bishop, C., Spiegelhalter, D. J., and Winn, J. (2003). "VIBES: A variational inference engine for Bayesian networks." In Becker, S., Thrun, S., and Overmayer, K. (eds.), *Advances in Neural Information Processing Systems*, volume 15, 777–784. Cambridge, MA, USA: MIT Press.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). "Latent dirichlet allocation." *Journal of Machine Learning Research*, 3: 993–1022.

Braun, M. and McAuliffe, J. (2007). "Variational inference for large-scale models of discrete choice."
  URL http://arxiv.org/abs/0712.2526.

Coburn, C. E. and Russell, J. L. (2008). "Distric Policy and Teachers' Social Networks." *Educational Evaluation and Policy Analysis*, 30(3): 203–235.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., and Harshman, R. (1990). "Indexing by Latent Semantic Analysis." *Journal of the American Society for Information Science*, 41(6): 391–407.

Fisher, D., Smith, M., and Welser, H. (2006). "You Are Who You Talk To: Detecting Roles in Usenet Newsgroups." In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, volume 3, 59b.

Johnson, C. M. (2001). "A survey of current research on online communities of practice." *Internet and Higher Education*, 4(1): 45–60.

Phan, X.-H., Nguyen, L.-M., and Horiguchi, S. (2008). "Learning to Classify Short and Sparse Text and Web with Hidden Topics from Large-scale Data Collections." In *The 17th International World Wide Web Conference (WWW 2008)*, 91–100. Beijing, China.
  URL http://gibbslda.sourceforge.net/

Reagans, R. and McEvily, B. (2003). "Network Structure and Knowledge Transfer: The Effects of Cohesion and Range." *Administrative Science Quarterly*, 48(2): 240–267.

Sawchuk, S. (2008). "Sites Mimicking Social Networks Set Up for Staff Development." *Education Week*, 28(1): 1–18.

Welser, H. T., Gleave, E., Fisher, D., and Smith, M. (2007). "Visualizing the signatures of social roles in online discussion groups." *Journal of Social Structure*, 8(2).

# Appendix

## A    Topics in the Forums

Table 2: Each of the 20 topics in the Forums identified by LDA is listed with a few of it's most informative high-probability words.

| Teachers, Tech & Professional Development | Classroom Tech | Tech & Subject Matter | Spam I |
|---|---|---|---|
| technology | students | history | create |
| learning | can | digital | lesson |
| teachers | online | art | customized |
| education | teachers | studies | smartboard |
| skills | classroom | social | new |
| development | content | world | resources |
| schools | access | pc | classroom |
| **Classroom Help Seeking** | **Technical Help Seeking** | **Collaboration** | **Foreign Language** |
| looking | has | students | language |
| ideas | anyone | project | english |
| some | software | will | session |
| thanks | school | school | spanish |
| class | open | group | public |
| using | used | collaboration | world |
| reading | experience | interested | esl |
| **Presentation Media** | **External Resources** | **Blogging** | **Spam II** |
| video | parents | blog | wordpress |
| powerpoint | student | please | learning |
| presentation | research | about | using |
| file | university | post | classroom |
| audio | assessment | students | interested |
| youtube | state | share | thanks |
| flash | program | comments | started |
| **Wiki's & Google** | **Workshop Announcements** | **Future of Education** | **Ed & Social Networking** |
| google | will | information | web |
| wiki | be | will | classroom |
| page | at | people | social |
| use | conference | change | tools |
| site | online | future | ning |
| search | time | education | education |
| rss | information | thinking | discussion |
| **Links** | **Foreign Words** | **School Vocab** | **General Vocab** |
| http | e | one | I |
| com | de | children | to |
| www | die | teacher | the |
| interesting | la | first | it |
| check | dem | kids | that |
| out | y | after | what |
| site | und | school | about |