© 2010 Eric Gilbert

# COMPUTING TIE STRENGTH

BY ERIC GILBERT

# DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science in the Graduate College of the University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

**Doctoral Committee:** 

Associate Professor Karrie Karahalios, Chair Associate Professor Brian Bailey Associate Professor ChengXiang Zhai Associate Professor Christian Sandvig Associate Professor Loren Terveen, University of Minnesota Dr. Jonathan Grudin, Microsoft Research

## ABSTRACT

Relationships make social media *social*. But, not all relationships are created equal. We have colleagues with whom we correspond intensely, but not deeply; we have childhood friends we consider close, even if we fell out of touch. Social media, however, treats everybody the same: someone is either a completely trusted friend or a total stranger, with little or nothing in between. In reality, relationships fall everywhere along this spectrum, a topic social science has investigated for decades under the name *tie strength*, a term for the strength of a relationship between two people. Despite many compelling findings along this line of research, social media does not incorporate tie strength or its lessons. Neither does most research on large-scale social phenomena. In social network analyses, a link either exists or not. Relationships have few properties of their own.

Simply put, we do not understand a basic property of relationships expressed online. This dissertation addresses this problem, merging the theories behind tie strength with the data from social media. I show how to reconstruct tie strength from digital traces in online social media, and how to apply it as a tool in design and analysis. Specifically, this dissertation makes three contributions. First, it offers a rich, high-accuracy and general way to reconstruct tie strength from digital traces, traces like recency and a message's emotional content. For example, the model can split users into strong and weak ties with nearly 89% accuracy. I argue that it also offers us a chance to rethink many of social media's most fundamental design elements. Next, I showcase an example of how we can redesign social media using tie strength: a Twitter application open to anyone on the internet which puts tie strength at the heart of its design. Through this application, called We Meddle, I show that the tie strength model generalizes to a new online community, and that it can solve real people's practical problems with social media. Finally, I demonstrate that modeling tie strength is an important new tool for analyzing large-scale social phenomena. Specifically, I show that real-life diffusion in online networks depends on tie strength (i.e., it depends on social relationships). As a body of work, diffusion studies make a big simplifying assumption: simple stochastic rules govern person-to-person transmission. How does a disease spread? With constant probability. How does a chain letter diffuse? As a branching process. I present a case where this simplifying assumption does not hold. The results challenge the macroscopic diffusion properties in today's literature, and they hint at a nest of complexity below a placid stochastic surface. It may be fair to see this dissertation as linking the online to the offline; that is, it connects the traces we leave in social media to how we feel about relationships in real life.

to Leslie & Carolyn

# ACKNOWLEDGMENTS

# TABLE OF CONTENTS

LIST OF TABLES	9
LIST OF FIGURES	12
CHAPTER 1: INTRODUCTION	17
What is Tie Strength?	19
Scope	19
Existing Approaches	20
Scenarios	20
Origin	21
Contributions	21
Fix a Problem; Get Some Data	23
Arc of this Dissertation	23
CHAPTER 2: LITERATURE REVIEW	25
Tie Strength	25
Framing other problems with tie strength	26
Modeling tie strength	29
This work	31
Socially-rendered Social Media	31
This work	34
Diffusion	41
Diffusion and tie strength	44
This work	45
CHAPTER 3: COMPUTING TIE STRENGTH	46
Research Questions	47
The Language and Architecture of Facebook	47
The Specific Goal of this Study	48
What does one number mean?	48
Method	49
Participants	49
Predictive Variables	50

Intensity variables	52
Intimacy variables	52
Duration variable	52
Reciprocal services variables	53
Structural variables	53
Emotional support variables	54
Social distance variables	54
Demographic and usage variables	54
Dependent Variables	54
Statistical Methods	55
Network effects	56
Standardization and normalization	57
Results	58
When the Model Breaks Down: High Residuals	65
Asymmetric Friendships	65
Confounding the Medium	66
Discussion	67
Individual predictors	68
Different kinds of tie strength	69
Conclusion	70
Theoretical implications	70
Practical implications	71
CHAPTER 4: DESIGNING AROUND TIE STRENGTH	73
Research Questions	73
The Language and Architecture of Twitter	74
The Collapsed Context Problem	76
We May As Well Call It "Temporal Media"	77
Twitter and Tie Strength	77
We Meddle	79
Origins of We Meddle	79
We Meddle Lists	80
We Meddle's Client	84
Communities	87

More meddling	87
Breaking the temporal model	88
Deployment	88
Architecture	90
Does It Generalize?	92
Mistakes in terms of predictors	93
How Users Experienced We Meddle	97
Reactions on Twitter and the Web	101
CHAPTER 5: INFORMATION DYNAMICS & TIE STRENGTH	107
Research Questions	108
Diffusion vs. cascades	109
Method	109
Retweets as the mark of diffusion	112
Results	112
Tie strength and content	117
Discussion	122
Note on the Onnela et al. graph	123
CHAPTER 6: LIMITATIONS, FUTURE WORK & CONCLUSIONS	124
Models and Predictors of Tie Strength	124
Alternate models of tie strength	125
Rendering Social Media Socially	126
Generalizability	128
Learning About Other Things With Tie Strength	128
Conclusions	129
APPENDIX A: LIST OF PREDICTIVE VARIABLES	131
APPENDIX B: TIE STRENGTH SEMI-STRUCTURED INTERVIEW PROTOCOL	134
APPENDIX C: WE MEDDLE SEMI-STRUCTURED INTERVIEW PROTOCOL	135
BIBLIOGRAPHY	136

# LIST OF TABLES

**Table 1.** Thirty-two of over seventy variables used to predict tie strength, collected for each of the 2,184 friendships in our dataset. The distributions accompanying each variable begin at zero and end at the adjacent maximum. Most variables are not normally distributed. The Predictive Variables subsection expands on some of these variables and presents those not included in this table.

**Table 2.** The five questions used to assess tie strength, accompanied by their distributions. The distributions present participant responses mapped onto a continuous 0-1 scale. This model predicts these responses as a function of the variables presented in Table 1.

**Table 3.** The intercorrelations of the five dependent variables. With the exception of *Job-Strong*, *Job-Loan* and *Bring-Job*, the dependent variables are well-correlated with one another.

**Table 4.** The top predictive variables as measured by standardized beta coefficients. The two Days since variables have large coefficients because of the difference between never communicating and communicating once. The utility distribution of the predictive variables forms a power-law distribution: with only these fifteen variables, the model has over half of the information it needs to predict tie strength. *F*-statistics reflect the impact of a particular feature, in this table and in each upcoming predictive variable table.

**Table 5.** The top predictive variables as measured by standardized beta coefficients for *How helpful would this person be if you were looking for a job?* Light gray rows indicate variables of particular interest.

**Table 6.** The top predictive variables as measured by standardized beta coefficients for *If you left Facebook for another social site, how important would it be to bring this person along?* Light gray rows indicate variables of particular interest.

**Table 7.** The top predictive variables as measured by standardized beta coefficients for *How would you feel asking this friend to loan you \$100 or more?* Light gray rows indicate variables of particular interest.

**Table 8.** The top predictive variables as measured by standardized beta coefficients for *How upset would you be if this person unfriended you?* Light gray rows indicate variables of particular interest.

**Table 9.** The top predictive variables as measured by standardized beta coefficients for the *How strong*? model in Chapter 3 and their Twitter analogs. We Meddle uses these features to build models of tie strength in Twitter. *Logged follower difference* substitutes for the original social distance variable, *Education difference*. It captures a difference in fame on Twitter. It's coefficient is decreased from the original because *Logged follower difference* is a substitute: the -0.123 is instead an average of all the social distance variables in Chapter 3. *Days since first communication* also factors in the ordinal value in which the user followed another user.

<sup>†</sup> The *Intimacy* category here is an agglomeration of many variables, an index. It blends two intimacy word measures, reciprocal following relationships, number of mutual following relationships and *Days since first communication*. *Structural* is *Median strength of mutual friends*, as in the original study.

# Both Structural variables refer to Median strength of mutual friends.

*§ Structural* again refers to *Median strength of mutual friends*, while *Social Distance* refers to the fame differential variable on line 6.

**Table 10.** An analysis of We Meddle's *strong tie* mistakes in terms of the model's input predictors. Also shown at the end of the table are non-model variables collected for various reasons: diagnostics, interface reasons or as the output of the model itself (tie strength). The table compares two groups, correct predictions and incorrect predictions, but limited only to those users who made a correction. + *median* refers to the median among the correct strong tie predictions; – *median* refers to the median among the correct strong tie predictions; – *median* refers to the median among incorrect strong tie predictions. W refers to the Wilcoxon statistic. Note the strong recency effect as indicated by *Days since last communication* and the many effects for network-based predictors: *Intimacy x Structural, Mean strength of mutual friends, Structural x Structural* and *Belongs to community. Belongs to community* here means that the partner (not the user) belonged to one of the biggest four communities generated for interface reasons. A Bonferroni correction for all the tests done here means that  $\alpha = 0.05/18 = 0.00278$ . (All "< 0.001" rows are therefore significant under standard hypothesis testing norms.)

**Table 11.** An analysis of We Meddle's *weak tie* mistakes in terms of the model's input predictors and non-model variables. The table compares two groups, correct predictions and incorrect predictions, but limited only to those users who made a correction. Table headings refers to same things as the table above. We see very different reasons for weak tie mistakes than we do for strong tie mistakes. The network-based predictors all but vanish, with *Belongs to community* contributing, but in exactly the opposite direction that it does for strong ties. We also see strong effects by *Logged follower difference* and *@-reply words exchanged*, perhaps signaling the ease with which you can message a higher-status user on Twitter. As opposed to the strong tie mistakes, we see that model makes mistakes as a function of tie strength, with mistakes more likely the lower the tie strength.

**Table 12.** The results of logistic regressions to predict automatically-inferred topics of retweeted URLs. Each regression controls for the same things, including fixed effects (modeled as dummy variables for each screen name). *LRT* refers to the  $X^2$  likelihood-ratio statistic. After controlling the family-wise error rate with a Bonferonni correction for the 17 topics ( $\alpha = 0.05/17 = 0.00294$ ), I can only confirm an effect (negative) of tie strength on *Politics*. However, tie strength also seems closely related to the *Entertainment & Culture* (positive) and Technology & Internet (negative) topics. Models where overall variance did not improve significantly are marked as "whole model rejected."

**Table 13.** The results of negative binomial regression on estimates of subjective language in retweeted URLs. I used the Opinion Finder system to automatically label both subjective sentences and direct subjective expressions in the text of the retweeted URLs. (Both terms are defined in the main text.) The regressions control for the same basket of variables as in the topic regressions, plus controls the sensitivity of the classifier: logged objective sentences and logged bytes. Without filter-

ing the data, tie strength exerts no influence on subjectivity. However, filtering out classifier noise on small documents does yield effects, though not strong ones and in opposite directions. One interpretation is that tie strength positively affects subjectivity but negatively affects automaticallylabeled sentences—meaning that it positively affects the amount of informal language.

# LIST OF FIGURES

**Figure 1.** Social media now and then. On the left, twitter.com renders a stream of status updates from people the user has followed, generating a social network. On the right, a user logs in through a terminal to read and write email with the program PINE. While we may be tempted to think of social media as a new internet fad, social media is as old as the internet itself.

**Figure 2.** The interface for SNARF, a social sorting email client. Reproduced from (Fisher et al. 2006).

**Figure 3.** A reproduction of the interface from Personal Map, a social representation of your email. Reproduced from (Farnham et al. 2003).

**Figure 4.** A reproduction of the interface from Contact Map, a social representation your contacts and messages. Reproduced from (Whittaker et al. 2002).

**Figure 5.** The Facebook Newsfeed and its reliance on time as the primary design element. Lately, it is clear that Facebook is ordering Newsfeed stories by some mix of time and other features, but from the outside it seems hard to tell exactly what.

**Figure 6.** With the exception of threading, email continues to rely on time as its main design element. In almost every fundamental respect, modern email clients look like the first email clients.

**Figure 7.** The social link-sharing site Reddit relies on time to present its comments, as does almost every blog-like site or technology (e.g., The Huffington Post, Scobelizer, etc.). Simple social tools (much simpler than understanding tie strength) seem like they could enhance time spent on sites like these, such as "only show me comments by people who've commented along with me in the past."

**Figure 8.** Internet Relay Chat (IRC), a group-chat technology with a long history, uses time to organize its design. It very much resembles Instant Messaging (IM), but participants' social bonds are often much more ephemeral.

Figure 9. A Usenet client from the 1980s using messages ordered by time to organize its design.

**Figure 10.** A reproduction of a simulated diffusion process based on tie strength, from (Onnela et al. 2006). The simulations show that if tie strength operates on the diffusion process (i.e., diffusion is a function of tie strength), we should expect very different macroscopic properties.

**Figure 11.** The questions used to assess tie strength, embedded into a friend's profile as participants experienced them. An automated script guided participants through a random subset of their Facebook friends. As participants answered each question by dragging a slider, the script collected data describing the friendship. The questions reflect a diversity of views on tie strength.

**Figure 12.** The model's Adjusted *R*<sup>2</sup> values for all five dependent variables, broken down by the model's three main terms (added in sequence from top). Modeling interactions between tie

strength categories results in a substantial performance boost. The model performs best on *Loan \$100*? and *How strong*?, the most general question.

**Figure 13.** The predictive power of the seven tie strength categories, presented here as part of the *How strong?* model. A category's weight is computed by summing the absolute values of the coefficients belonging to it. The diagram also lists the top three predictive variables for each category. On average, the model predicts tie strength within one-tenth of its true value on a continuous 0–1 scale.

**Figure 14.** The model's performance across all ties in the dataset. There is a strong correlation, with Adj.  $R^2$  above 0.5. Yet the model shows a very slight bias toward underestimation, represented as the larger cloud in the bottom-right of the figure. The gap in the center results from participants' inclination to move the slider from its starting point, if only slightly.

**Figure 15.** The data from the previous figure, recast as a binary classification problem. *Weak tie* is interpreted as those ties participants rank as below average, and *Strong tie* as those above the mean. (x-axis  $\mu$ =0.43 and y-axis  $\mu$ =0.46.) Under this simple reformulation, the model performs at roughly 89% classification accuracy.

**Figure 16.** The popular microblogging service Twitter (http://twitter.com) in 2010. Users of the service "follow" others users in a asymmetric way (i.e., reciprocation not necessary) and their status updates appear all together in a stream such as this one. The stream allows for passive social awareness. From a consumption perspective, the collapsed context problem means that people from various parts of your life are all squished together in one place, the single stream.

**Figure 17.** The popular Tweetdeck client offers users another way to interact with Twitter. Roughly 10% of Twitter users use Tweetdeck; at least twice as many use twitter.com. It is, however, based on the same principles as the Twitter web interface. Tweetdeck has a columnar layout, with most columns replicating existing Twitter pages, such as "Mentions." The leftmost column, for example, is essentially the same as twitter.com. We Meddle Lists, introduced shortly, can plug into existing tools like Tweetdeck.

**Figure 18.** The homepage of http://wemeddle.com. This is the screen a user sees before she grants We Meddle access to her Twitter credentials. The page gives the user an opportunity to watch a demo video I made for new users by clicking the "demo video" link in the bottom left corner. The video is (currently) available at http://wemeddle.com/wemeddle.mov.

**Figure 19.** The We Meddle wait screen during a user's first login. Clouds float by on screen as We Meddle digs through the user's Twitter history. The text below says "We're chatting with the ghost of Twitter past. It might take a minute. He's old, you know." After the computation completes (mostly I/O bound), the user gets redirected to We Meddle Lists, illustrated below.

**Figure 20.** The We Meddle Lists web interface, grouping a user's Twitter contacts with an underlying computational model of tie strength. Upon sign-in, We Meddle automatically computes tie strength for everyone the user follows. The two groups on the left (e.g., Inner Circle and Outer Circle) result from the tie strength computation. The two groups on the right come from community detection on the underlying social network. (See text for details on its construction.) **Figure 21.** The author's inner circle, but this time I have dropped two accounts from the automatic suggestions before I make the list. We Meddle suggests people who belong as strong ties, but invites user make corrections, under the assumption that it's easier to subtract a few accounts rather than add each one individually.

**Figure 22.** A We Meddle Lists user who left feedback about the site on We Meddle's accompanying online community. She's happy with the concept, but wonders why We Meddle does not allow list additions. I explain this decision in the text, but in short: I constrained the interface to better ensure I could study tie strength.

**Figure 23.** The We Meddle Client web interface, rendering a user's Twitter stream using an underlying computational model of tie strength. Here, a user has moved the slider slightly toward the inner circle to emphasize strong ties at the expense of weaker ones. We Meddle translates emphasis into profile image size, typeface, font size and opacity. Users can also apply filters (upper right) to reduce the stream. For example, with the interface still rendered by tie strength, users can filter the stream to only particular social communities, like web researchers.

**Figure 24.** The We Meddle Client filtering tweets to only show ones it thinks are positive. (This is a different stream than the one above.) We Meddle uses a set of word-matching heuristics borrowed from previous literature to mark tweets as positive or negative. See text for details.

**Figure 25.** A heatmap of We Meddle's user distribution across the world. Most users come from the United States, but China also has a strong presence, accounting for between 6% and 12% of users (depending on whether you measure in visits or raw page views). The maximum on the legend in the lower-left corner is 2,403. (Image generated by Google Analytics.)

**Figure 26.** The architecture of We Meddle. When a user signs in for the first time, a process forks off hundreds and sometimes thousands of requests against the Twitter API. As they come back, they filter through the simple application of the tie strength model and then percolate through the user's ego network. Usually in under one minute, We Meddle generates its lists from the output of the model.

**Figure 27.** A screenshot from an impromptu reaction to an We Meddle a talk by Prof. Karahalios. The talk was webcast live from the Berkman Center for Internet and Society at Harvard. I captured this portion of the talk and made it available at http://social.cs.uiuc.edu/berkman.we.meddle.mov.

**Figure 28.** An unprompted reaction from a Twitter user. The fact that We Meddle recreated lists he had already made for himself speaks to computational tie strength's face validity.

Figure 29. Berkman fellow David Weinberger reacts to We Meddle.

Figure 30. A We Meddle user comments on his lists during Prof. Karahalios's Berkman talk.

**Figure 31.** We Meddle has received many visits from non-English speakers. This tweet roughly translates as: This tool gets your Twitter friends (!?), processes them and afterward gives you lists. When you see the results you can edit however you want. ^\_^

**Figure 32.** A Romanian user writes: Oh geez, this application is a great filter for your Twitter contacts.

**Figure 33.** A Chinese user writes: This is a tweet from this link [client link]. We Meddle is pretty good, it automatically divides my friends into groups, so that occasional tweet from my important friends won't be buried in an ocean of other tweets.

**Figure 34.** Another Chinese user writes: This thing is miraculous. I want to know how it works very much. It even classifies my GF's account into Inner Circle, but I never send tweets to her account, and it also classifies accounts I don't care about into outer circle. The 4 categories are very accurate, almost 100%. It is miraculous.

**Figure 35.** We Meddle has started conversations, as well as acting as a way to sort a user's own contacts. In this tweet, a We Meddle user asks the Facebook researcher Cameron Marlow if he participates in the online community MetaFilter. (A "MeFite" is a MetaFilter member.)

**Figure 36.** In this blog post, a user of We Meddle screencapped the We Meddle interface and commented on it. The user goes on to tell a story about the people in the Inner Circle, even apologizing for those left out.

**Figure 37.** A retweet using the "built-in" mechanism for reposting tweets. Twitter provides a button (and a formal API mechanism) to redistribute someone else's tweets to your followers. In this case, the user dabbish and five others have retweeted physorg\_com's tweet, which contains a link to the full story.

**Figure 38.** A retweet using a user-adopted syntax, in this case the via syntax. Here, idsgn has retweeted GOOD's tweet about a video. The syntax retweet predates the built-in retweet, and many users still prefer it. This style retains a user's identify: syntax retweets appear to come from the retweeter, not the retweeted, as above.

**Figure 39.** A high-level overview of the method for collecting and preparing data for the analysis presented in this chapter. After capturing three types of retweets and preprocessing, I generate the three main datasets I analyze in this chapter: multimedia retweets, topic-coded retweets and subjectivity-coded retweets.

**Figure 40.** The reference distribution of all We Meddle tie strengths (gray) and the distribution of tie strengths from retweets (filled yellow). If tie strength had no effect on retweeting behavior, the two would closely align. Instead, this graph shows that tie strength affects retweeting behavior, with stronger ties bulging out at near 0.6 and 1. A Kruskal-Wallis test confirms this visual intuition. (Note: these curves represent probability distributions, smoothed with a Gaussian kernel. Each curve is generated from 4,158 data points.)

**Figure 41.** The reference distribution of all We Meddle tie strengths (gray) and the distribution of tie strengths from retweets containing links to photos, videos and music (filled brown). Retweets with multimedia links look similar to retweets generally, but skews to stronger ties. Post-hoc Wilcoxon tests confirm the difference between media retweets and retweets. (Note: these curves repre-

sent probability distributions, smoothed with a Gaussian kernel. Each curve is generated from 4,158 data points.)

**Figure 42.** The *unscaled* pseudo-CDF of retweet probability conditioned on tie strength. This curve uses Bayes's Rule to turn the two figures above into  $P(RT | ts \le t)$ . This graph is essentially the ratio of the filled graph in Figure 40 to the gray line. It is unscaled because this curve leaves out P(RT), the baseline probability of retweet. However, as a constant (e.g., it's median across users) probably suffices for P(RT), this curve will drive overall estimates of P(RT | ts  $\le$  t).

The curve illustrates how extremely weak ties get retweeted more often than expected. (The smoothing process in the previous figures obliterated it.) With the exception of extremely weak ties, weak ties generally suffer in retweets. This curve shows lots of detail due to its 0.01 step size. However, we see a somewhat steady rise between 0.2 and 0.6, with a jump near 0.6 and a slower rise afterward. Extremely strong ties (e.g., near 1) get a big bump too.

**Figure 43.** Another estimate of the unscaled pseudo-CDF of retweet probability conditioned on tie strength, this time with a larger step size. The step size in this figure is 0.02, whereas the previous figure uses 0.01. Notice how the effect for extremely weak ties goes away, as that one data point  $[P(RT \mid ts \le 0.01)]$  has been smoothed in a larger bin. The overall trends remain the same. This curve provides more robust estimates, but also sacrifices detail and interesting discontinuities.

The red line shows a Multivariate Additive Regression Splines (MARS) fit of the data. In this case, MARS learns a simplified piecewise linear approximation of the curve. It segments the data into three parts, each with distinct slopes: segments I will call *weak ties, medium ties* and *strong ties*. Based on the high variability we see at the endpoints of Figure 42, we might add the segments *very weak ties* and *very strong ties*.

**Figure 44.** The method for estimating P(RT) from 635 We Meddle users. Over a one week period, I obtained the size of each user's tweet stream, the number of tweets that would have come through their client had they been logged in the entire time. Counting the number of retweeting relative to the overall size of the stream gives you an estimate of P(RT). It's not the whole story, however, since users may be logged out or simply not paying attention. It's debatable whether this should factored in or not, as I have done here.

**Figure 45.** The probability distribution of P(RT) at user level, from 635 We Meddle users. The median of this distribution is 0.0001262704, but I collapse all activity into one estimate of P(RT) = 0.0001426229.

**Figure 46.** A illustrative triad clipped from a broader social network where one node is strongly tied to two others, with tie strengths 0.92 and 0.87. Ideas from structural balance indicate that suggesting a completed triad is probably a bad idea.

# CHAPTER 1 INTRODUCTION

When we chat via email or instant messages, do we leave clues about the closeness of our relationship? If so, what clues? How often we talk? How often I initiate the conversation, or how often you do? How quickly we reply to one another? The particular words and phrases we type to each other? Our positions in our social networks? The purpose of this dissertation is to find these answers, and to show that they matter for the design and analysis of social media.

In modern social media like Facebook, Twitter and email, relationships are the stuff that makes the medium *social*. However, take a look through your email address book or Facebook friend list. Reflect on your relationships with the people there. Before long, I bet you will agree that not all relationships are created equal. We have colleagues with whom we correspond intensely, but not deeply; we have childhood friends we consider close, even if we fell out of touch. Or, take this example reported in the press: some Human Resources departments have taken to cold-calling an applicant's Facebook friends instead of asking for references! One HR manager said that by using social media "you've opened up your rolodex for the whole world to see." Of course, sometimes they call someone hoping for references, "only to find that you were just drinking buddies" (Tahmincioglu 2008).

Academics see it too. For decades, various social sciences have documented how different types of relationships affect individuals and organizations (Granovetter 1983). In this line of research, relationships are measured in the currency of *tie strength* (Granovetter 1973). Loose acquaintances, known as *weak ties*, can help a friend generate creative ideas (Burt 2004) or find a job (Granovetter 1974). They also expedite the transfer of knowledge across workgroups (Hansen 1999). Trusted friends and family, called *strong ties*, can affect emotional health (Schaefer et al. 1981) and often join together to lead organizations through times of crisis (Krackhardt & Stern 1988). Despite many compelling findings along this line of research, social media does not incorporate tie strength or its lessons. Instead, all users are the same: friend or stranger, with little or nothing in between. Most empirical work examining largescale social phenomena follows suit. A link between actors either exists or not, with the relationship having few properties of its own (Adamic & Adar 2003; Albert & Barabási 2002; Nowell & Kleinberg 2003). We simply do not understand a fundamental property of how relationships express themselves online. Consider the following quote from a special issue of PNAS on computational social science, a reflection on what actually constitutes a social relationship:

... Important questions remain unanswered. Observers have debated whether digital communications offer new methods of creating intimacy or are inflated measures of social connectedness that skim the surface of real attachments. Despite these important issues, little is known about whether electronic data indeed are a valid proxy for the real social connections they purportedly measure. Previous work has not scientifically addressed the level of agreement between actual social ties and electronic communication means. — "What is a Social Tie?" (Wuchty 2009)

In other words, tie strength is a blind spot. This dissertation addresses this problem, merging the theories behind tie strength with the data from social media. I show how to reconstruct tie strength from digital traces in online social media, and how to apply it as a tool in design and analysis. Tie strength is more than a methodological or theoretical preoccupation; a model of tie strength has the potential to significantly impact social media users. Consider automatically allowing the friends of strong ties to access your information on a site, without having to set any permissions ahead of time. Or, as one of my participants cleverly suggested, consider remaking Facebook's Newsfeed to get rid of "people from high school I don't give a crap about."



Figure 1. Social media now and then. On the left, twitter.com renders a stream of status updates from people the user has followed, generating a social network. On the right, a user logs in through a terminal to read and write email with the program PINE. While we may be tempted to think of social media as a new internet fad, social media is as old as the internet itself.

Sometimes, when we talk about social media, it's easy to think of it as a new internet fad. However, social media is old, as old as the internet perhaps (Figure 1). It's important to review this point at the outset. Consider what this 1977 article had to say about the role of email in the early days of the internet:

The initial goals in creating the ARPAnet were to promote more effective use of geographically dispersed computing facilities ... A new use emerged, however ... network message service was an immediate success. Message flow grew in volume to become the most visible (if not the heaviest) traffic on the network. (Henderson Jr. & Myer 1977)

Even then, at the beginning of the internet, email (a *social medium*) had a tremendous presence. Today, social media has exploded on the internet: Facebook boasts over 500 million users and is one of the most visited sites on the internet. Yet, both Facebook and email have roots in human language and personal relationships which build over time. Although the sheer scale is different, many things remain the same. We still type to each other. We still build up relationships, and generate and articulate social

networks. But still questions remain. Do relationships express themselves online in predictable ways? Can we automatically infer what they mean? If so, can it help us build and analyze social media?

## WHAT IS TIE STRENGTH?

This dissertation is about tie strength. Tie strength is a diffuse, and sometimes analytically unsatisfying concept. It refers to a general sense of closeness to another person. When that feeling is strong, we call it a *strong tie*; when it is weak, we call it a *weak tie*. Who are you close to? Who are merely your acquaintances? These kinds of questions commonly pop up in tie strength studies. Tie strength does not reference a particular situation—although we see it applied as analytic tool in many different ones, such as the analysis of business transactions. Mark Granovetter, who introduced the concept, had only this to say about the question: "Most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent." I adopt a similar outlook here.

This rather general, diffuse viewpoint on tie strength has advantages. Say we encounter a new context we have not studied yet. Imagine we have studied instant messaging in the workplace, but not how non-profits use email to coordinate. We could start from scratch. But why not use what we know about the general concept of tie strength in human relationships as a starting point? This principle underpins the work I present here: learn about a general and important attribute of human relationships in order to use it in new, uncharted contexts.

#### SCOPE

Tie strength, of course, can mean closeness in real life or in some mediated channel, like email or Facebook. Many of us maintain relationships in online media as often (or sometimes even more often) than in real life. Most existing tie strength work is set in real life (Granovetter 1983). Questions like "How do mixtures of strong and weak ties affect someone's ability to succeed in the workplace?" commonly pop up in these studies.

The focus of this dissertation, however, is how tie strength expresses itself in social media on the internet. While I study the ways in which tie strength shows itself through traces in social media, the dependent variable is the very "real life" question, "How strong is your relationship with this person?" (In follow-up interviews, it was clear that participants interpreted the question in its real-life sense.) Clearly, not everyone from our real life comes with us online. As I have presented this work in various forms, inevitable questions arise: "My mom does not email me. We always call each other. How can you model that?" I discuss these situations in more detail in Chapter 3. At a high level, however, it may be fair to see this dissertation as linking the online to the offline; that is, linking social media traces to how we experience relationships in the real world.

#### EXISTING APPROACHES

I'm not the first person to care about tie strength, or the first person to try to model it. Until now, however, we have used heuristics to estimate tie strength. I take a different approach here. This dissertation is the first to show that you can reconstruct tie strength meaningfully from a basket of traces left in a social medium.

In most tie strength work, tie strength itself is not the object of interest. For instance, all of the following have substituted for tie strength at one time or another: communication reciprocity (Friedkin 1980), possessing at least one mutual friend (Shi et al. 2007), recency of communication (Lin et al. 1978) and interaction frequency (Gilbert et al. 2008; Granovetter 1973). Instead of studying tie strength itself, these studies wanted to examine macroscopic network properties, or the effect of relationships on job hunting. Tie strength is only a tool.

Is a heuristic like "call it a strong tie if they message each other at least N times" good enough? If it's bad, how bad? Before this work, we did not know. However, from my data I can now estimate that this commonly used heuristic classifies strong vs. weak ties at roughly 61% accuracy (letting N = 10 for this example). A fuller model, like the one I built from Facebook data in Chapter 3, classifies strong vs. weak ties at roughly 89% accuracy. (In other words, different people communicate different ways at different times. Frequency does not usually work as a substitute.) Some recent work has shed light on how much we distort findings by relying on relational heuristics. In *WWW 2010* paper, De Choudhury et al. show that the macroscopic properties of large networks are very sensitive to how you define ties between two nodes. This work dovetails nicely with what I present here: this dissertation offers a way to construct weighted ties from social media data, something fundamental to how we represent data in computational social science.

#### SCENARIOS

If we could model tie strength, how could we actually use it? Could we build something around it? I demonstrate one potential application in this dissertation (a re-rendering of social streams via tie strength), but we can imagine many. To see what I mean, consider a woman interacting with her friends and family via a social network site. She posts photos, talks about her job, her family life and how night classes are coming along. And she uses it to keep up on everybody else's life. Now let's imagine that the next time she logs in, she's been on vacation for a month with limited internet access. She wants to catch up. What should the system show her? Everything? Probably not. Ideally, we would show her the most important things that happened to the most important people. Maybe her best friend changed jobs, or her sister took some great photos of *her* vacation. Understanding the core of personal relationships is the first step to building systems that can do these things.

Other scenarios come to mind, as well. Most social media applications on the internet allow users to set privacy levels, policies like "allow this person access to photos, but not this person," "permit this person to post content on my profile, but not these others," and so on. And, they are notoriously hard to manage. In fact, the blog TechCrunch recently quoted<sup>1</sup> Mark Zuckerberg, the founder of Facebook, as saying "Guess what? Nobody wants to make Lists," referring to Facebook's affordance for grouping friends in lists and applying policies en masse. Note that he uses the word "make," not the word "use." Two things make it a particularly thorny problem: 1) social relationships and boundaries are often fuzzy and change over time; 2) it is simply a lot of work to manually sort everyone you know into groups. Tie strength can make this problem tractable. A system that understands tie strength might make reasonable initial guesses at who gets what access, which users could subsequently clean up. We could make a lot of headway getting users 90% of the way.

#### ORIGIN

Before we go much further, I feel compelled to explain the origins of this dissertation. It does not follow the trajectory of a typical Computer Science dissertation, or even a Human-Computer Interaction one. There was no vexing interface problem; no users crying out for a tie strength model. This dissertation started with a simple question after reading Granovetter (Granovetter 1973) for the n<sup>th</sup> time: What is tie strength *exactly*? Surely, I would find the answer by the end of the day, in a paper I had missed. Surprised, I did not. I would find it tomorrow. Nope.

Whereas most HCI and Computer Science dissertations begin with an irritating interface problem or how you pump more percentage points out of a computational model, this one did not. It started with a simple question about the nature of personal relationships in online media. Only afterward did I consider the problems we might now attack with such a model—and they pop up everywhere. It became clear that new theoretical problems also become tractable with a metric for tie strength. In this dissertations, I present explorations in these two directions.

#### CONTRIBUTIONS

This dissertation makes the following specific contributions:

1. A rich, high-accuracy and general way to reconstruct tie strength from digital traces.

The model presented in Chapter 3 uses more than 70 carefully-chosen, theoreticallymeaningful indicators of tie strength. It is built from Facebook data, and it splits users into strong and weak ties with nearly 89% accuracy. Among many potential applications, the model offers a computational way to rethink and redesign social streams, the topic I explore with We Meddle in Chapter 4. With We Meddle, I show that the model presented in Chapter 3 general-

<sup>&</sup>lt;sup>1</sup> http://techcrunch.com/2010/08/26/facebook-friend-lists

izes to a new online community in which the model did not train: Twitter. It was unclear that it would. Maybe dynamics differ so much between the two communities that a general purpose model does not work. However, Chapter 4 reveals an error structure just like the Facebook model, evidence for its generalizability. This is the first work to demonstrate stable interpersonal relational properties across online media, a new direction in online communities research (Kittur & Kraut 2010). Chapter 4 also examines the model's mistakes in terms of its predictors, revealing directions for future refinements.

#### 2. A social media application on the web which puts tie strength at the heart of its design.

I built an application for Twitter users called *We Meddle*, open to anyone on the web with a Twitter account. It applies the tie strength model presented in Chapter 3 to a user's contacts and interaction history in Twitter, a different social medium than the one in which the model trained. We Meddle is both an experimental platform and simply a tool I hope makes social media a little bit better. It is the first application I am aware of to put a calibrated relational model at the heart of its design. With it, I examine both the generalizability of computational tie strength and its worth in design. Overall, feedback from users has been very positive and suggests that computing tie strength can solve real people's real problems. Over 1,300 people from around the world have used We Meddle with no coercion or payment; they used it because they thought they would find value in it. I report on its design, its architecture, the reaction it received on the web and on follow-up interviews with users.

# 3. Findings showing that real-life diffusion is a function of tie strength. That is, real-life diffusion in online networks depends on social relationships.

As a body of work, diffusion studies make a big simplifying assumption: simple stochastic processes govern person-to-person transmission (Fowler & Christakis 2008; Gruhl et al. 2004; Kempe et al. 2003; Kossinets et al. 2008; Liben-Nowell & Kleinberg 2008). For example, how does a disease spread? With constant probability across the network. How does a chain letter diffuse? As a branching process. I present a case study where the simple stochastic assumption does not hold. I study the spread of links across the Twitter network and show that transmission is a function of tie strength. This is the first large-scale, quantitative work I am aware to show that tie strength affects real-life diffusion practices. Moreover, I explore the subtle interaction between tie strength and content. For example, the results show that political content diffuses less frequently through weak ties, and that tie strength also operates on subjective content, but in strange ways. These findings suggest very different macroscopic properties than what the literature suggests (Onnela et al. 2008), and they hint at a nest of complexity below a placid stochastic surface.

#### FIX A PROBLEM; GET SOME DATA

In part, this dissertation explores where computing tie strength can improve how users experience social media today, a classic HCI approach. At the same time, I want to collect enough data to test the the ideas behind computational tie strength, something we might call computational social science (Lazer et al. 2009). Although computational social science is clearly still emerging, leaders of the field have expressed concerns that it remains at the mercy of the corporations with the data: large telecoms, internet service providers, big social media sites, etc. How can the field move forward when the data lives behind a locked door? Many companies do not engage with the research community. The group of distinguished researchers in (Lazer et al. 2009) suggests solutions like a national research clearinghouse for computational social science data, with clear ethical guidelines and open access.

However, as a subtext in this dissertation, I suggest another way. Fix a problem; get some data. The Twitter client I present in this dissertation, *We Meddle*, attacks the *collapsed context problem*, the collapse of our social contexts within online media. At the same time, I collect valuable data on how tie strength works in a new medium. I did not pay, beg or coerce my users to come to We Meddle. They came out of curiosity and because they thought We Meddle might help them. As a result, I have 145,000 Twitter relationships annotated for tie strength and, in some cases, corrected by users. From this data, we can do social science.

Of course, researchers cannot hope to compete against the mainstream products of corporate giants like Google, Yahoo or Microsoft. I would not recommend building a competitor to GMail to do social science. You will get killed. However, for many, many reasons, holes exist in these products. As another example, I recently completed a side project called *Link Different*<sup>2</sup>. From a social computing perspective, it covers well-trodden ground: Link Different fixes a visibility problem in Twitter by telling you how many of your followers have seen a link before you broadcast it. 1,300 people have used We Meddle, but well over 50,000 people have used Link Different. As a consequence, I have data on threshold and conformity effects in information diffusion: When do people choose to share as a function of the interest of their audience? When is it okay to be like everybody else? 50,000 data points is not the 500 million we could get from Facebook's entire database, but it's quite a bit better than zero.

This is little more than a subtext to this dissertation. However, it's one that I hope excites researchers: do not wait for someone to present you with data on one knee. Fix a problem; get some data.

#### ARC OF THIS DISSERTATION

Next, I will review a large body of literature on tie strength, socially-rendered interfaces and diffusion. (However, Chapters 3, 4 and 5 also present a good deal of related work in situ, hopefully making it eas-

<sup>&</sup>lt;sup>2</sup> http://linkdifferent.net

ier to compare and contrast it with my work.) In Chapter 3, I lay out the backbone of the dissertation: a computational model of tie strength (Gilbert & Karahalios 2009). The model incorporates theoretically meaningful variables from the social sciences and performs with high accuracy. In Chapter 4, I present *We Meddle*, a web-based application for Twitter users that puts tie strength at the heart of its design. Through it, I examine the generalizability of computational tie strength and also its value in design. In Chapter 5, I study tie strength in a forgotten area of diffusion research: the actual mechanism of person-to-person transmission. I study if and specifically how tie strength regulates the flow of information across the Twitter network, possibly interacting with content along the way. In the final chapter, I revisit the limitations of my work and where future work could go.

# CHAPTER 2 LITERATURE REVIEW

Tie strength is one of the most influential concepts in sociology. In the twenty-seven years since its publication, the paper introducing tie strength (Granovetter 1973) has attracted over 15,000 citations<sup>3</sup> from a wide range of fields, like Organizational Studies, Finance and Computer Science. It would be impossible to cover them all. Instead, this chapter focuses on the most seminal, relevant and provoca-tive papers from the tie strength literature.

My work models tie strength from traces in social media and applies the model as a design and analytic tool. In this chapter, I review these three main topics: model, design, analyze. Correspondingly, Chapters 3, 4 and 5 find their foundations in the literature presented here. First, I consider tie strength in the literature of many social sciences. Next, I explore the (somewhat scant) work on socially-rendered interfaces, drawing upon research in Human-Computer Interaction. Finally, I visit the diffusion literature, work that examines how facts, opinions and even physical things flow through networks. At the end of each section, I bring the contributions of this dissertation into relief, showing how this work extends what we know about each field.

### TIE STRENGTH

In this section, I consider the tie strength literature. After reviewing seminal tie strength work, I examine research that places tie strength front and center as an analytic frame, but whose main concern is something else (e.g., communication, economic inequality, the spread of rumors, etc.) I conclude this tie strength review by visiting work that models tie strength, and present a lens through which to interpret this dissertation's contributions.

Mark Granovetter introduced the concept of tie strength in his landmark 1973 paper "The Strength of Weak Ties" (Granovetter 1973). (In the rest of this chapter, I will refer to "The Strength of Weak Ties" as SWT, in the interest of compactness.) In it, he defined tie strength as follows:

The strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie. (Granovetter 1973)

Granovetter left the precise definition of tie strength to future work. He did, however, richly characterize two types of ties: *strong* and *weak*. Strong ties are the people you really trust, people whose social circles tightly overlap with your own. Often, they are also the people most like you. (i.e., homophily). The young, highly educated and metropolitan tend to have diverse networks of strong ties (Marsden

<sup>&</sup>lt;sup>3</sup> http://scholar.google.com/scholar?q=strength+of+weak+ties

1987). Weak ties, conversely, are merely acquaintances. Weak ties often provide access to novel information, information not circulating in the closely knit network of strong ties. Granovetter develops his tie strength framework in the context of people hunting for jobs (a consistent theme throughout Granovetter's work).

In SWT, Granovetter paints tie strength in broad strokes. The paper, while hugely influential, is exploratory. At times, he seems to be commenting on sociology itself. For context, at the time of its publication most sociology analyzed networks composed of strong ties. In this research tradition, strong ties hold communities together, whereas weak ties fracture communities. In this article, Granovetter asserts rather the opposite: weak ties provide the bridges that make communities viable over the long term. SWT repeatedly links strong ties to dense networks, what some call a cluster: dense communities inside larger networks. While using broad strokes for the majority of the article, Granovetter does speculate (but only speculate) on the specific structure of tie strength the construct, "Should tie strength be developed as a continuous variable?" I will revisit this conjecture in my own modeling work.

In a 1983 follow-up article (Granovetter 1983), Granovetter revisited his very influential argument with a round-up of studies that adopted tie strength, using it as their analytic lens. For instance, Granovetter reviews research demonstrating that scientific discoveries flow more efficiently through weak ties than through strong ones, a not-so-intuitive finding. In another instance, Lin et al. (1978) asked participants to deliver a booklet to some unknown person in a distant place. (The Lin experiment is a more principled recreation of a classic Milgram experiment (Milgram 1967).) Lin et al. examined the rate of success and the characteristics of the paths the booklet took, including a measure of tie strength operationalized by two features, recency of communication and relationship type (i.e., "friend" or "acquaintance"). They found that people who used more weak ties in their paths had more success reaching their destination. This is just one among many examples validating the strength of weak ties argument. However, Granovetter also writes, "Lest readers of SWT and the present study ditch all their close friends and set out to construct large networks of acquaintances, I had better say that strong ties can also have value." Strong ties provide emotional support, are more stable and easier to rely upon. Organizations rely on them during difficult times (Krackhardt & Stern 1988). Granovetter cites many studies showing that the poor rely heavily on dense networks of strong ties. In a modern article (Granovetter 2005), Granovetter shows how strong ties in particular affect macroeconomic outcomes.

## Framing other problems with tie strength

Many researchers have adopted tie strength as an analytic lens for studying their own problems (Granovetter 1983). Here I review only some of these findings (because there are so many), and I put particular emphasis on ones relevant to this dissertation or just particularly stunning ones. Strong ties provide social support that can actually improve mental health (Schaefer et al. 1981), something we might expect since values (e.g., religion, thriftiness) flow along strong ties. We see this reinforced by Fowler and Chrisakis's recent work: in a long-term study, they document how happiness moves along social ties (Fowler & Christakis 2008). In this work, happiness (the emotional variable of interest) is shown to transfer to when people are surrounded by many other happy people. In other words, Fowler and Christakis present a threshold model where associating with lots of happy people makes you happy. While they do not explicitly frame their work in terms of tie strength (they do not even cite Granovetter), this threshold effect model echoes the arguments Granovetter makes in SWT strong ties, represented by dense ties in the threshold model, can transfer emotion to their connected peers.

Researchers have found traction applying tie strength to organizations, as well. For instance, banks that find the right mix of weak and strong ties with other firms tend to get better financial deals (Uzzi 1999). The finding is impressive and worth quoting:

Firms are more likely to get loans and to receive lower interest rates on loans if their network of bank ties has a mix of embedded ties and arm's-length ties. These network effects arise because embedded ties motivate network partners to share private resources, while arm's-length ties facilitate access to public information on market prices and loan opportunities so that the benefits of different types of ties are optimized within one network.

"Embedded" here is a synonym for dense clusters of interconnected nodes, again associated with strong ties. Note that this study concerns ties between *firms* not individuals, a key mark of this and related organizations work (Adler & Kwon 2002; Reagans & Zuckerman 2001). Tie strength scales beyond strictly personal relationships to relationships of groups of people, each one representing a firm. Note that at an operational level, we again see a reliance on a single measure of tie strength, "embeddedness" (i.e., network density), acting as a proxy for tie strength.

As I alluded earlier, strong ties between employees from different organizational subunits can help an organization withstand times of crisis (Krackhardt & Stern 1988). Yet, strongly tied coworkers are also the ones likely to create crises by pushing for institutional change (Krackhardt 1993). In these two pieces, Krackhardt examines the role of social relationships inside firms experiencing change, from within and from without. The idea is that social networks form within structured contexts, according to physical and interactional proximity (i.e., friends often become friends by mere exposure). Those relationships can either help or hurt the firm in times of change. Just as in other related work (Levin & Cross 2004), Krackhardt finds that different types of ties serve different roles at different times. Building on this theme, and in perhaps the most well-known extension of the SWT argument, Ronald Burt has advanced structural holes theory (Burt 1995). In its original formulation, Granovetter's SWT argument makes everyone sound passive: you're one node in a giant network, sometimes seeking, but

often merely exposed to information. Interestingly, diffusion work adopts a similar perspective: the agents are passive and rarely act in their own interest. Burt's work is different. Here, people embedded in networks use their positions for their own gain. Someone who acts a bridge (in the network sense; also known as a weak tie) has the ability to control, shape and regulate information flowing between groups. Sometimes he can benefit from the position. For example, imagine a developer who eats lunch with people from R&D. Over lunch one day, they tell our developer about their new research application and their findings; early knowledge of it puts the developer's group in position to make a strong bid when the company decides to develop it as a product. This is how filling a structural hole confers advantage, both to the group and to the person acting as the bridge. In a followup study on this theme (Burt 2004), Burt shows how people who occupy the bridge role often get credited with having more creative ideas. In this study of an American electronics company, Burt shows how people situated at the boundary between subgroups have higher "compensation, positive performance evaluations, promotions, and good ideas."

Having covered how tie strength affects emotion, health, finance, and organizations, I conclude with tie strength's place in Computer-Mediated Communication (CMC). Up to this point, we have not considered mediated relationships. What I have presented so far lives in the world of face-to-face relationships: we see each other every day, have coffee every so often, or chat at the office. But to this dissertation, CMC matters even more. (To frame this review, note that early work implicated the internet as a destroyer of social cohesion and prosocial behavior (Kraut et al. 1998), while later consensus found quite the opposite (Wellman et al. 1996).) Perhaps the most well-known result from the CMC tie strength literature finds a mediated analog to the Granovetter hypothesis: weak ties also act as conduits for useful information in CMC (Constant et al. 1996). Set in a larger organizational context (a U.S. tech company), Contstant et al. find that weak ties will provide technical assistance to one another even when they have little or no past relational history. This is a subtle, but important point. In immediate (i.e., face-to-face) contexts, weak ties generally mean acquaintances. Here, in a CMC context, a weak tie means something more like a familiar stranger (Milgram 1977): someone who inhabits the same space as you (e.g., forums, chat rooms, etc.), but with whom you have no relationship in real life, and barely one even in mediated life. Still, Constant et al. found that broad institutional goals (i.e., the company's well-being) translated into offers of help to relative strangers. We see similar claims in recent research on question-and-answer sites (Hsieh 2009). The Constant et al. finding is an important backdrop to this dissertation: it frames which online activities correlate with which kinds of ties, but leaves unanswered how to specifically represent and reason about relationships online.

Another core concept from CMC and tie strength (and one especially relevant to this dissertation), is *media multiplexity*. In short, media multiplexity means that in mediated relationships, I can chose

among many possible channels when I talk to you (e.g., email, phone, IM, social media sites, etc.). We simply do not have this multiplexity in real life; it's a feature unique to CMC. The key finding on media multiplexity comes from Haythornthwaite: weak ties tend to use a handful of commonly available media, while strong ties often chat through many channels (Haythornthwaite 2002; Haythornthwaite & Wellman 1998). This finding, while very well-represented in the literature since its publication (see (Williams et al. 2006) and (Hampton & Wellman 2003) for representative publications), has attracted a recent challenge (Hogan 2009). In his Ph.D. dissertation, Hogan argues from survey data that media multiplexity has an unexplored structural dimension: multiplexity interacts with network structure, not only tie strength. For now, we can say that media multiplexity is a core result in CMC and tie strength, but one that may need further study or renovation. It's important to remember the multiplexity findings when interpreting what I present in this dissertation. I explore only a single medium at a time, and so multiplexity suggests that my method misses something key about how strong ties interact online. I revisit this limitation in Chapters 3 and 6.

### Modeling tie strength

Even in his own work, Granovetter uses how many times someone *saw* someone else as a proxy for tie strength (Granovetter 1974). Granovetter explains in footnote that tie strength the concept had not occurred to him when he did the research, but this practice of finding proxies has continued into the present. In theory, tie strength has many manifestations. In practice, relatively simple proxies have substituted for it: communication reciprocity (Friedkin 1980), having a mutual friend (Shi et al. 2007), recency of communication (Lin et al. 1978) and interaction frequency (Gilbert et al. 2008; Granovetter 1973). As David Krackhardt put it:

At what point is a tie to be considered weak? This is not simply a question for the methodologically curious ... the theory makes a curvilinear prediction. How do we know where we are on this theoretical curve? Do all four indicators count equally toward tie strength? (Krackhardt 1993)

Krackhardt published that work twenty years after SWT. So, two questions remain. First, the simple one: How do we actually measure tie strength? Second, the more complex: What *is* tie strength exactly? As I quoted above, Granovetter proposed four tie strength dimensions: amount of time, intimacy, intensity and reciprocal services. But he did not lay out to how to actually build tie strength. Subsequent research has only added to Granovetter's original four categories. Burt (as described more fully above) proposed that structural factors shape tie strength, factors like network topology and informal social circles (Burt 1995). Wellman and Wortley argued that providing emotional support indicates a stronger tie (Wellman & Wortley 1990). In their view, offering advice on family problems, for instance, would indicate a stronger relationship. Lin et al. (1981) show that social class, embodied by factors like socioeconomic status, education, political affiliation, race and gender, influences tie strength. They all seem like plausible and potentially important indicators.

Perhaps the closest existing work to put them all together is Marsden and Campbell's 1984 paper "Measuring Tie Strength," which I will discuss in detail as it so closely bears on this dissertation. Marsden and Campbell open the paper by saying, "Little attention has been given to the measurement of the concept of tie strength." (It seems somehow strange that even 11 years after the wildly successful SWT, no one had tried to nail it down.) In their paper, Marsden and Campbell use survey data from three cities (two American and one German) to estimate the weights on various predictors of tie strength. "Closeness," the major dependent variable of interest, was measured trichotomously in their survey: "acquaintance," "good friend" or "very good friend." But the survey also included the tie strength dependent variables frequency, duration, breadth of discussion and mutual confiding. The survey asked participants to recall their three closest friends and ten data points describing those friendships. The predictors used in their analysis included neighbor status, coworker status, overlapping network measures, prestige difference (in occupation) and educational difference. They make the following observations: time and depth describe tie strength well; "closeness" is the best measure of tie strength; frequency and duration are problematic as predictors. This is a very important paper in the development of tie strength, and I will continually compare my results to its findings. But I would also like to point out some of its weakness, many of which Marsden and Campbell note themselves. It is built on retrospective data from informants, and therefore likely contains exaggerations and underreporting (Bernard et al. 1984; Marin 2004; Marsden 1990). Their design only asks participants about their three closest friends, yielding results clustered at the strong end. Marsden and Campbell did foundational work. I am, however, claiming that this dissertation fills in these gaps, and at the same time describes how tie strength works in CMC.

Until now, I have featured work coming from the social sciences. Quite recently, however, computer scientists like myself have worked on tie strength (Easley & Kleinberg 2010; Kahanda & Neville 2009; Viswanath et al. 2009; Xiang et al. 2010). In their very recent networks textbook, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*, Easley and Kleinberg devote an entire chapter to tie strength. However, the book focuses on *networks*, and the chapter treats tie strength structurally: tie strength effectively equals the shortest path length between two vertices in the graph. In a more relational approach, (Kahanda & Neville 2009) formulates tie strength as a link prediction problem, a classification task where you have to predict a "top friend" link. (Some Facebook users installed an application allowing them to mark their "top friends.") Xiang et al. (2010) focus on a latent graphical model of tie strength, where tie strength is operationalized as common photos or common wall posts. In a different approach, Viswanath et al. (2009) study relational activity in time, finding that ties usually

"decay," meaning that we see less and less communication over a tie as time goes on. (This is perhaps to be expected, as new friendships often require lots of early maintenance; however, that does not necessarily imply tie strength gets correspondingly weaker). In all these papers, the focus is on the computational model and on learning it efficiently. It seems quite natural to a computer scientist to frame tie strength this way: How *hard* is it? Is it *harder* than *Problem X*? As I argue shortly, I think this is the wrong formulation. Instead, the power lies in the features we elicit.

### This work

This dissertation makes new contributions to this line of literature, primarily in Chapter 3. My work leans heavily on social science, finding much of its early inspiration there. Specifically, the work presented in Chapter 3 shows how to model tie strength in online, mediated spaces. It also presents a transparent model. We can easily look inside the model to see how to build tie strength. This is a new chapter in what we understand about tie strength, very much in the tradition of Marsden and Campbell. But particularly for CMC and tie strength, this dissertation constitutes a big step for how we understand online relationships at perhaps their most basic level. What kinds of communicative acts actually mean something in online spaces? Does recency matter more than frequency? How much does language matter? How much does time (and time decay) matter? This work answers those questions.

When forming my basket of tie strength predictors, I channel lots of social science literature. This stands in stark contrast to the computer science papers I just discussed. The focus there is on the model and learning it efficiently, not features. I argue that these approaches do not capture tie strength because they do not elicit it from the population. In other words, people communicate with one another for all sorts of reasons in all sorts of contexts—adopting any one metric recalls the early proxies of the 1970s and 1980s. By comparison, the model I present here is simple by design, but mixes many metrics together. It derives its power not from model-engineering, but from the richness of the features them-selves. I argue that this is a profound strength: the simplicity of model makes it generalizable. In Chapter 4, I discuss how this model of tie strength generalizes past the specific online community that generated it. Without this simplicity, I suspect generalization would have been impossible—as most case studies of domain adaptation in the machine learning literature suggest.

## SOCIALLY-RENDERED SOCIAL MEDIA

After modeling tie strength from traces left in social media, in Chapter 4 I try to render a social media interface by tie strength. Can a social media interface actually reflect the social relationships represented formed in it? This may seem like a banal question: of course they can! They already do! But take a close look at social media interfaces. Nearly universally, they rely on time to organize their interfaces, as their central design axes. I offer the following series of screenshots, Figures 5 through 9 as evidence for this claim. Although it's been discussed for quite a while, few social systems have attempted to place

an actual, important relational construct at the heart their design. As a result, almost all social media is designed around when messages appear. Figures 5 through 9 all make the same underlying point: every user is the same, trusted friend or total stranger. (Yet, as some simple hacks have shown (Lieberman & Miller 2007), social renderings can have very positive outcomes.)

However, three projects have tried. SNARF (Fisher et al. 2006; Neustaedter et al. 2005), Personal Map (Farnham et al. 2003) and ContactMap (Whittaker et al. 2004) all explore using interaction histories to reconfigure interfaces. However, it may be fair to describe them as exploratory or feasibility studies. As these three projects represent, more or less, the scope of the literature on socially-rendered interfaces, I will discuss them in detail. (It is perhaps important to note that some social media interfaces like Apple's Mail.app offer search relevance scores that may adopt some social data into the final score. Also, the Facebook Newsfeed seems to prioritize information by some metric. However, academics have a hard time comparing their systems to these commercial systems because the commercial ones are closed. We do not know how they work.)

SNARF, the "Social Network and Relationship Finder," is a social sorting prototype designed to solve the email overload problem. Too much mail comes from too many sources: bosses, spouses, family members, long lost friends, chatty colleagues, mailing lists and so on. SNARF uses features from past email exchanges to visually depict which people are most important to a SNARF user. The features include "emails sent to each person from the user," "replies to each person from the user," "emails CC'd to each person from the user," "emails to the user and marked unread," and so on. In total, SNARF extracts 11 features. You can then select any one of them as the key upon which the system sorts your email, as depicted in Figure 2. In Figure 2, differently sized and differently saturated bars backfill each person who has sent email to the user. The main idea is that the interface uses history to help people visually triage their email. In a follow-up study, the authors report that users found SNARF useful when under time pressure, but that SNARF did not help their self-reported perceptions of email overload (Fisher et al. 2006). In Personal Map (Farnham et al. 2003), the authors aim to give people a way to group their contacts according to social groups, providing a meaningful grouping of their address book chaos. Again, the tool applies heuristics (undocumented in the paper) to infer rough groups of people in the user's email life. As depicted in Figure 3, Personal Map then visualizes the results for you. The authors also built a add-on to Personal Map that auto-suggests likely candidates for addressees on a email given some initial seed names. The authors report that users in a study generally found the tool useful and enjoyable.

Contact Map (Whittaker et al. 2004; Whittaker et al. 2002) explores a similar idea. In it, the authors develop hypotheses about who is most likely to be "important" and perform a logistic regression to predict it from six features. However, users in a lab study did not respond well to these automatic scores. The authors use this features along with something akin to community detection to create a "social desktop," an alternate way to interact with organizational social media, like email. Figure 4 is a reproduction of its interface. Contact Map is early inspiration for We Meddle, and it bears some resemblance to the work I conduct in this disserta-

tion. The authors fully develop the idea of a "social desktop," a representation of interpersonal communication that centers around people instead of around messages.

Outside of HCI, perhaps the closest body of computer science work is in recommender systems. My work employs similar techniques. However, it's very important to note a big difference:

a recommender system almost always connects people with information or products. Consider Netflix's movie recommendations or Amazon's "Customers Who Bought This Item Also Bought" feature. One core principle underlying all these systems is attribute-based social homophily: I am more likely to enjoy movies people like me enjoy, where similarity is measured in terms of age, prior viewing habits, gender, etc. Exemplified by (Garg & Weber 2008) and (Zanardi & Capra 2008), to take only a few examples from the ACM conference on Recommender systems, recommender systems often live in systems that are not primarily social. You do not go to Netflix just to interact with other people, certainly not your friends; you go there to watch movies. The focus in social media is simply different and it re-

🛎 SNARF	
Menu	
Unread To/CC me	<b>a</b>
Darcy Adler (3)	
Cynthia Allen (2)	
Tim Andrews (2)	
Unread Mail	🚍 🖾
Darcy Adler (3)	^
Chris Adler (1)	-
Cynthia Allen (2)	
Tim Andrews (2)	
Angela Nowak (2)	
Kris Andrews (1)	
Otto McCoy (4)	
Betty Esposito (1)	
Samantha Turner (1)	
Carl Hilton (3)	
Kara Esposito (2)	×
This Week's Mail	<b>a</b> 🛛
Steve Nichols (14)	^
Mona Silva (19)	
Preethi Ramos (3)	
David Ingram (19)	
Tom Alvarez (2)	-
Dylan Andrews (7)	
Dorris Andrews (7)	
Drew Alvarez (2)	
Dorris Allen (136)	
Chandra Ortega (83)	
Missy Sewell (12)	~
Last E-mail change: 12:43 PM	11

Figure 2. The interface for SNARF, a social sorting email client. Reproduced from (Fisher et al. 2006).

mains an open problem to find the connection between recommender systems and tie strength, but not one I consider in this dissertation.

## This work

This dissertation is a new step in this (admittedly short) line of work on socially-rendered social media. We Meddle, presented in Chapter 4, is an extension of the ideas in SNARF and Contact Map. It extends existing work in three ways. First, it renders social streams in the wild on the internet. Second, it explores using tie strength as the backbone for interface techniques like auto-lists and social zooming (demanding visual attention as a function of tie strength). Third, We Meddle is the first application I know of to put a *calibrated* relational model at the heart of its design. In fact, one way to see the present work is a test of whether users find value in (or even will tolerate) a rich, entirely computational tie strength system. (Users did not respond well to automatic "importance" scores in Contact Map (Whittaker et al. 2004).) Whereas earlier systems lived in the lab, anyone on the internet can use We Meddle. This is a step forward for the field. These are not internal corporate users with a limited set of objectives and lots of common ground. It's the internet.

Aside from the academic literature reviewed above, it also makes sense to compare this work to applications available commercially. We Meddle calculates tie strength automatically in the background, without user action. This is important. Some current interfaces allow users to group their friends or the accounts they follow (e.g., TweetDeck's groups<sup>4</sup>, Facebook's Friend Lists), but users need to put in lots of effort to build them. Anecdotal evidence in the press from Facebook executives, referenced earlier, suggests that users rarely use these features.) We Meddle wants users to do as little work as possible.

<sup>4</sup> http://tweetdeck.com



Figure 3. A reproduction of the interface from Personal Map, a social representation of your email. Reproduced from (Farnham et al. 2003).



Figure 4. A reproduction of the interface from Contact Map, a social representation your contacts and messages. Reproduced from (Whittaker et al. 2002).


Figure 5. The Facebook Newsfeed and its reliance on time as the primary design element. Lately, it is clear that Facebook is ordering Newsfeed stories by some mix of time and other features, but from the outside it seems hard to tell exactly what.

000				Sen	t (2892	messages)	
							Q
MAILBOXES	§ •	То	Subject	Date Sent	v	Size	Subjects Computing Tip Strength - Edday 10:00 pm LIS 126
/ Inhox		Karahalios, Kyrat	Re: I have received permission to	1:15 PM		1.2 KB	From: Eric Gibert
Deafter (TD)		Karrie G. Karahal	100 people logged in during your talk	12:50 PM		0.9 кв 🗸	Date: March 2, 2010 11:19:08 AM CST
Toraits (15)	•	Karahalios, Kyrat	Can you reset my wemeddle accou	11:22 AM		15.0 KB	To: kham@uluc.edu
-n∕ Sent	•	twidale@illinois.e	Computing Tie Strength - Friday 10	11:20 AM		4.1 KB	
🕨 🗑 Trash 🛛 😰		twidale@illinois	Computing Tie Strength - Friday 10	11:20 AM		0.7 KB	Hi Kevin, I'm giving a practice job talk at GSLIS on Friday. I would love it if you could make
V ON MY MAC		kham@uiuc.edu	Computing Tie Strength - Friday 10	11:19 AM		1.7 KB	it. It's at GSLIS because I'm interviewing at information schools.
info		hci@cs.uiuc.edu	Computing Tie Strength - Friday 10	11:11 AM		1.7 KB	Cheers, Eric
werneddle	•	mheuer@medsou	н	11:10 AM		6.4 KB	Computing Tie Strength
		social@cs.uiuc.edu	Karrie's (webcast) talk at Berkman	11:01 AM		0.6 KB	Computing the Strength
Kecover	•	Karrie G. Karahal	Have you made interview decisions	10:58 AM		14.5 KB	Relationships are the heart of social media: they make it *social*.
▶ CS 🕢	•	Inobert Floyd	Practice talk	10:54 AM		3.3 KB	Yet from a computational standpoint, we understand very little, if
		Inobert Floyd	Re: Practice talk	10:54 AM		2.7 KB	for instance, what makes two people close? Is it mostly a function of
		Inobert Floyd	Practice talk	Yesterday		0.6 KB	how long they"ve been friends? How often they talk? What they say? Or
		Inobert Floyd	Re: Reserving Room 126 on March 5	10:53 AM		6.8 KB	relationship as tie strength, the difference between a best friend and
	•	Joshua Hailpern	commencing attack on Andrew	10:50 AM		4.6 KB	a distant colleague. I will present my work on modeling tie strength
	•	sarita.vardi@oma	feedback	Yesterday		13.1 KB	computationally, and how it can act as a tool in both design and
	•	Joshua Hailnern	[social] Do we have old machines si	Yesterday		2.0 KB	a Twitter application I built that puts tie strength at the heart of
	•	Forrest Jandola	[cs-sophomores mailing list] Check	Yesterday		9 3 KB	its design. We Meddle explores both the generalizability and design
	•	Karabalios, Kyrat	30 logins from cs-undergrads	Yesterday		9.4 KB	research agenda comprising exciting problems facing social computing.
		Balamuta, John W	Re: [social] Online Lynch Moh?	Yesterday		1.2 KB	such as how ultra-successful communities can guide new contributions
		Kevin Sperry	Re: Roger Ebert	Yesterday		1.0 KB	by learning from old ones.
		brant.chee@gmai	Banding together	Yesterday		2.0 KB	
	•	Sonhy Yi	At a different number	Yesterday		2.2 KB	
		Karabalios, Kyrat	Re: Michael	Yesterday		1.4 KB	
	•	Dan Cosley	Interview date	Yesterday		14.1 KB	
		Karabalios, Kyrat	Re: [social] Group Meeting tomorro	Feb 28, 2010	n	2.3 KB	
	•	Cliff Lamne	Question about Michigan	Feb 28, 2010	n	4.2 KB	
	•	enilber2@cs.uiuc	test	Feb 27, 2010	5 D	1.9 KB	
		Loren Terveen	Re: Important CSCW issue: Should	Feb 27, 2010	- n	18.1 KB	
		Karrie G. Karahal	Check it out!!	Feb 27, 2010	n	0.9 KB	
		Kare Anderson	Re: meddle not working right now	Feb 27, 2010	D	1.7 KB	
	•	Leslie Gilbert	Your visit to Michigan	Feb 27, 2010	0	22.1 KB	
		Karrie G. Karabal	Words of wisdom	Feb 27, 2010	- D	0.8 KB	
	•	Lada Adamic	I hear you're my host	Feb 27, 2010	0	5.3 KB	
	•	Kerr, Joanne	Generic logistical email and items f	Feb 26, 2010	- D	40.6 KB	
	•	Sandvig, Christia	New We Meddle client	Feb 26, 2010	- D	2.0 KB	
		Ingbert Floyd	Talk	Feb 26, 2010	D	1.7 KB	
	•	mheuer@medsou	Saturday	Feb 26, 2010	D	11.5 KB	
		Paul Lambert	CS undergrads and update	Feb 26, 2010	D	0.7 KB	
	•	Karahalios, Kyrat	Cs undergrads	Feb 26, 2010	0	9.6 KB	
	•	social@cs.uluc.edu	[social] We Meddle Client social test	Feb 26, 2010	D	4.6 KB	
		Karrie G. Karahal	fvi: Cornell interview March 17	Feb 26, 2010	D	0.5 KB	
+ 🔺 🌣 -	•	Karrio C. Karabal	Stat undate	Fob. 26, 2010	- D	CE KP	

Figure 6. With the exception of threading, email continues to rely on time as its main design element. In almost every fundamental respect, modern email clients look like the first email clients.

PICS - R	EDDIT.COM - POLITICS - ASKREDDIT - FUNNY - WTF - GAMING - WORLDNEWS - PROGRAMMING - SCIENCE - TECHNOLOGY - ATHEISM - COMICS - IAMA - VIDEO	S - TODAYILEARNED - OFFBEAT - MUSIC - ECONOMIC! MORE >>		
(Ú) A (	SK REDDIT			
0	AskRedDIT comments related shirt	want to join? register in seconds   English		
113	Hey Reddit, I'm a meme, what should I do about this? (self.AskReddit) submitted 3 hours ago by IamAameme	search reddit		
*	I will not tell you what meme, but it consists of being in a picture on multiple websites. It isn't bad but over 30 people have recognized me and complete strangers. Any advice?	this post was submitted on 02 Mar 2010 <b>113 points</b> (67% like it)		
	EDIT 1: It turns out that I am really a meme, I'm approval guy. It's a pretty cool meme to be.	221 up votes 108 down votes		
	My friends and I were at an underwear party and I refused to get in my underwear. So my buddy started to take pictures and decided it would be hilarious to post them on 4chan. Two years later, it has taken off.			
	Here is the original & an additinal photo from the party as proof that I am really him.			
	Here it is on Dramatica and Know Your Meme	remember me recover password		
	It recently took off because it was on failblog	AckPoddit		
	I know one thing for sure, I approve of this meme.	Askitedult		
	225 comments share	- frontpage 127,927 people who should be		
		working.		
con	hments	For your questions only, please.		
sorte	ed by: <u>best</u> ▼ 200 ▼	GUIDE		
1 N	afarious- 131 points 3 hours ago [-]	<ul> <li>AskReddit is for thought-provoking.</li> </ul>		
* N	o one give him advice until he shows us the meme.	inspired questions.		
p	rmalink	<ul> <li>Please submit text-posts only (no outside links.)</li> </ul>		
-	IamAameme [S] 41 points 1 hour ago [-]	<ul> <li>Please don't post poll-type questions.</li> </ul>		
1	Ok, I updated it, now give me advice. SEE EDIT 1	<ul> <li>If you see spam, report it! It really helps.</li> <li>If you thick your past has disappoared.</li> </ul>		
	permalink parent	not hesitate to contact one of the		
	v0-z 18 points 1 hour ago [-]	mods, we're happy to help.		
	* Have you collected your internet dollars yet?	<ul> <li>Please be aware this sub is not a substitute for real legal or medical help.</li> </ul>		
	permalink parent	Subreddits to note:		
	KittyCatMan 5 points 1 hour ago [-]			
	What about the free internet coupons?	Relationship Advice     Self		
	permalink parent	• IAmA		
	Kozmo199 2 points 29 minutes ago [-]	Does Anyone Else     NeedAdvice		
	He's in line right behind sneezing panda.			
	permalink parent	Help HAITI - 100% of your donation will be used.		
	sodoh 2 points 31 minutes ago [-]			
	$^{lat}$ He has to wait until they pay the chocolate rain guy. Think he is first in line.	Ever read the reddiquette? Take a peek!		
	permalink parent	a community for 2 years		
	Raziel66 1 point 19 minutes ago [-]	MODERATORS message the moderators		
	Tr. Pepper paid him	noname99		
	permalink parent	tuber		

Figure 7. The social link-sharing site Reddit relies on time to present its comments, as does almost every blog-like site or technology (e.g., The Huffington Post, Scobelizer, etc.). Simple social tools (much simpler than understanding tie strength) seem like they could enhance time spent on sites like these, such as "only show me comments by people who've commented along with me in the past."

	#ubuntu I	#gentoo	I	##windows	I	#Freenode
@agli	(19:56:13) Joined the	#gentoo channel.				
@Alfred	(19:56:13) Channel 1	opic: Gentoo Linux suppor	t   Can't spe	ak? /join #gento	oo-ops   Paste	e over 4 lines?
@astinus @auoni	http://dpaste.com	Gentoo chans: xrl.us/kftd	KDE4? xrl.	us/otdxr + #gen	too-kde   e2fs	progs block fix:
@blackace	xrl.us/bea7ut I X S	erver 1.5? Bust KB/Mouse	? xrl.us/bem	6c4 1.6? http://	xrl.us/bfarit   p	oppler blocks?
@bonsaikitte	http://xrl.us/bephtt	Gentoo 10.0 LiveDVD: htt	to://xrl.us/aei	ntoo10vears I F	Profiles: 2008	0->10.0
@Naamah	nvidia-drivers/ glib	/ notl downgrade? sync an	d retry (bug	#286961)		
ere +aballier	(19:56:13) Tonic set	by rei on Tue Oct 06 16:37	57 2009			
+ABCD	(19:56:14) ** NOTIC	=: #windows: Forwarding to	another cha	nnel: ##windov	MS	
+Caster	(19:56:18) [marduk]	mrmowali: call?	vanourier ene	annet. ## anneoa	15	
+cia +dabbott	(19:56:29) [ marduk ]	: vou mean the interpreter?				
+darkside_	(19:56:32) [ marduk ]	i ); run(oxocuto				
+Dave2	(19:56:36) [ miniowg					
+doc home +Evetation	(19:56:41) [ hintitowy	n j. van n li opokiomonEtor: voob	that was wha	tiwaa oonfuqo	d about atk	and nautilus and
+Fieldy	(19.00.41) [ Diakkrief	n j. cookiemonoter, yean	unal was wha	ii rwas comuse	iu about. yik+	anu nautitus anu
+Griffon26	(19:56:49) [marduk]	proposed: which python?				
+Halcyon +hlieberman	(19.50.46) [ maruuk ] (10:56:52) [ marua outo	i himowyii. which pythons				
+idlor	(19:56:52) [ mmowg (19:56:59) [ mmowg	li j. yali li j: Looo				
+jayne	(19.56.56) [ mmowg	ii j. i see				
+Johnm_ +JonathanD	(19:57:08) [ mmowg	li j: got it I: 1: Theorem				
+kernelsense	(19.57.12)[mmowg	II J. I MANKS!				
+Levia	(19:57:49) [ marouk ]	: mrmowgli: use of vertical	space is goo	ba for coairig	not so good it	princ;) villavinalaaa
+LoKez +Martinpoo	(19:57:57) [ peedub ]	: anyone know what kernel	drivers are r	leeded for a m	x-gouuls rosev	viii wireless
+MinnieBann	adapter (pci)?					
+mpagano_	(19:58:35) ** Mad3 h	as quit [ "Leaving" "				
+mueli +nalioth	(19:59:02) ** RypPn	nas quit   "It would be so ea	asy, if I had t	ne right horse!"		
+neurogeek	(19:59:28) [ Skunky ]	peedub: what's ispci say it	tis?			
+niko	(20:00:53) ** Steven	M has joined the channel				
	(20:01:21) < Steven_	M is now known as: Guest	69819 >			
+robertz	(20:01:42) ** Guest6	9819 has quit   Remote clo	sed the con	nection **		
+RichiH	(20:02:03) [ peedub ]	: Skunky: says its the one i	have enable	d but still not w	orking was ju	ust wondering if any
+SpanKY +sping	one had one of the	ese working right under gen	too i had it	working on a L	FS install but o	an't get it in
+stew	gentoo					
+swegener	(20:02:32) [ Skunky ]	peedub: might be missing	ı firmware, d	epending on wh	hat chipset it u	ses.
+t	$\rightarrow$					

Figure 8. Internet Relay Chat (IRC), a group-chat technology with a long history, uses time to organize its design. It very much resembles Instant Messaging (IM), but participants' social bonds are often much more ephemeral.

X Her	mes	
X	6990 comp. rang. crarron	Min all cases [Minor point: In face of hard links there is no
	847 comp.lang.clipper	amonical name). This would suggest ention 1
	9969 COMP.LANG.CLIPPER	calonical namej. Inis would suggest option 1.
	4 comp.lang.clos	While option 2 it wields an IMMO more expected behavior when TRUENAME
	3 comp.lang.clu	is applied to any a stroom
	3898 comp.lang.cobol	13 appillou co say a scream.
	29 comp.lang.dylan	So we wight want both behaviors
	240 comp.lang.eiffel	30 we might wate both behaviors.
	1257 comp.lang.forth	Further
_	8 COMP. LANG. FORTH	
	1865 comp.lang.fortran	> function OST FOLLOW-LINK nathspec => nathname
	674 comp.lang.functional	> Function 051.Follow-link paulopec => paulitaile
	8 comp.lang.icon	Seturn the nathrane pointed to but the sumbolic link designated by
	15 comp.lang.idl	> PATHSPEC Signals a FILE-FEROR if PATHSPEC is wild or does not
	583 comp.lang.idl-pvwave	A designate a symbolic link
	16227 COMP.LANG.JAVA	accigitado a cynocizio inite.
	4933 comp.lang.javascript	Sunnose I do [unix nathname syntax assumed]
	2966 comp.lang.labview	bappeter i de faith palanane effetil decaneerj.
»	comp.lang.lisp	(ensure-directories-exists #o"/foodir/")
	52 COMP.LANG.LISP	(osi make-link #o"/foodir/" #o"/zoodir/")
	185 comp.lang.logo	(
7	118 comp. Lang. misc	How am I now supposed to follow the created link?
$\Delta$	robocup or bust!	11
»	operating system interface proposal	(osi:follow-link #p"/zoodir") ;A file, eh?
	lisp-2 or lisp-1	or
	robocup or bust!	(osi:follow-link #p"/zoodir/") ;A dir, hmm?
	operating system interface proposal	· · ·
	operating system interface proposal	E Further what do the following yield?
	operating system interface proposal	
	lisp-2 or lisp-1	(osi:probe-directory #p"/noodir")
	robocup or bust!	(osi:probe-directory #p"/zoodir")
	for people who use linguistic arguments	(osi:probe-directory #p"/zoodir/")
	lisp-2 or lisp-1	(osi:probe-directory #p"/foodir")
	seattle area lisp users' group meeting	(osi:probe-directory #p"/foodir/")
	lisp-2 or lisp-1	(osi:probe-directory #p"/foodir/bar.text")
	lisp-2 or lisp-1	
	for people who use linguistic arguments	Why is OSI:DIRECTORY-LIST needed, when we have CL:DIRECTORY?
	lisp-2 or lisp-1	
	robocup or bust!	Just a few nitpicking questions,
	lisp-2 or lisp-1	
	lopstr 2003 (call for abstracts)	Gilbert Baumann
5		;; Will hack for food.
1		N.

Figure 9. A Usenet client from the 1980s using messages ordered by time to organize its design.

# DIFFUSION

In the final piece of my dissertation, I study the role of tie strength in diffusion through Twitter networks. In a seminal review of diffusion research, Strang and Soule defined "diffusion" as follows:

Diffusion refers to the spread of something within a social system. The key term here is "spread," and it should be taken viscerally (as far as one's constructionism permits) to denote flow or movement from a source to an adopter, paradigmatically via communication and influence ... Diffusion is the most general and abstract term we have for this sort of process, embracing contagion, mimicry, social learning, organized dissemination, and other family members. (Strang & Soule 1998)

As "network science" has gained prominence and traction over the last 10 to 15 years (Benkler 2006; Borgatti et al. 2009; Cohen et al. 2000; Hidalgo et al. 2007; Jackson 2008; Potterat et al. 2002; Proulx et al. 2005; Watts 2004), diffusion research has exploded. How and why do people adopt new products? By what mechanism does a chain letter spread? What is the likelihood of an epidemic given an infection in a small part of a connected graph (some literature adopts the biological word "contagion" instead of "diffusion")? Diffusion research has attacked all these problems. Here, in this final part of the literature review, I will trace some of the seminal papers and trends in diffusion research, while at the same time carving out space for my own work on how tie strength modulates the flow of information in an online network.

Two classic, early studies of diffusion are Ryan and Gross's (1944) study of farmers adopting new seeds and Coleman, Katz, and Menzel's study (1957) of doctors adopting new treatments. Ryan and Gross document how farmers came to know about new, hybrid corn via salesman, but only planted it after they heard about success stories from nearby farmers. Coleman, Katz and Menzel study how doctors in three Midwestern cities came to hear about, and subsequently use, an risky new drug in their treatment regimens. Like Ryan and Gross, they cite "ongoing social practices," what we might today call "structural features," as influencing adopters. In other words, doctors were likelier to adopt the new drug after they heard about it from a colleague they knew socially. It's possible to interpret both of these studies at the interpersonal level, but they also speak to macro, birds-eye effects. These multiple granularities have continued to the present, with modern studies, such as one focused on the transmission of HIV (Rothenberg et al. 1998), highlighting the role of interpersonal dynamics and global network structure. However, it is important to note that (especially in light of the studies that come next) diffusion does not only happen via internal, interpersonal mechanisms. We often hear about new ideas, opinions and products from external sources like the media. While most research in this chapter focuses on social network diffusion, Rogers (1995) and Strang & Soule (1998) each present compelling evidence for the impact of external forces on diffusion patterns (Katz et al. 2006). We should interpret all diffusion research with this in mind.

Many researchers have advanced macro-level models of diffusion, In other words, what microscopic processes generate the macro-level data we observe in the world? Granovetter proposed a threshold model, where people adopt the thing diffusing after a certain number of their ties do (Granovetter 1978). If this model holds, then weak ties find themselves particularly disadvantaged, as threshold models preference a network's strongly connected clusters (i.e., strong ties), as Macy shows through simulations (Centola & Macy 2007) and Siegel shows through empiricism (Siegel 2009). However, other researchers question whether this model faithfully represents reality: they fail to find evidence for threshold models after searching for it in real world data (Strang 1996). The threshold model remains in limbo. Whereas sociologists have largely studied how structural properties impact diffusion, social psychologists have examined how proximity and power relationships influence diffusion processes (Festinger et al. 1963). (The finding on proximity affecting influence among dyads will be familiar to an HCI audience as the "distance matters" hypothesis (Olson & Olson 2000).)

The recent literature provides some very compelling case studies of diffusion. Often, researchers are interested in how something becomes normative or attractive. In a 2007 paper, Leskovec, Adamic and Huberman study the effect of person-to-person recommendations on products in an ecommerce web-

site. They find that while certain communities and types of products seem susceptible to viral marketing (i.e., inducing person-to-person recommendations), most are not. As with most work on "big data" diffusion, the authors posit "a simple stochastic model" for explaining macro-level behavior. While superficially different, others have studied romantic relationships and likelihood of disease transmission among teens (Bearman et al. 2004). Bearman et al. find that real-life romantic relationships do not carry some of the same structural features thought to promote the transmission of sexually transmitted diseases. A series of later empirical studies largely confirmed their simulated results (Klovdahl 2005).

Fowler and Christakis have conducted a series of studies at the intersection of diffusion and health (Christakis & Fowler 2007; Christakis & Fowler 2008; Fowler & Christakis 2008). All of the papers use the Framingham Heart Study dataset, a continuously running interview study from 1971 to 2003. (A relatively stable core of individuals remained present throughout the course of the study.) The three studies examine smoking, obesity and happiness. In the smoking study (Christakis & Fowler 2008), Christakis and Fowler show how whole clusters of the network seemed to quit smoking at nearly the same time, suggesting a strong effect for tied participants. They found even stronger effects for smoking cessation by a spouse, a sibling and a friend. Interestingly, they also found an effect for education: more educated peers influenced participants more than those with less education. In the obesity study (Christakis & Fowler 2007), Christakis and Fowler found similar effects, noting that obese people tend to influence each other's weight. They go to pains to demonstrate the effect independent of homophily. (This is a subtle, but crucial point. If two people become friends because of some shared trait or interest, we cannot attribute the diffusion to the tie. Christakis and Fowler controlled for it using repeatedly sampled longitudinal data.) They also note a same-sex effect: same sex dyads influenced each other more than opposite-sex dyads. Finally, in the happiness study (Fowler & Christakis 2008), Fowler and Christakis again control for the temporal confound, finding that "longitudinal statistical models suggest that clusters of happiness result from the spread of happiness and not just a tendency for people to associate with similar individuals." Geographical proximity also appeared as a significant predictor. Taken together, the Christakis and Fowler studies show that unexpected things (i.e., smoking, obesity and happiness) travel through social networks.

The diffusion concept spans many academic disciplines, traveling all the way to finance. Risk perception (e.g., tolerance for risky stocks in a portfolio) is linked to your social network (Scherer & Cho 2003). In particular, the authors find that "social linkages in communities may play an important role in focusing risk perceptions." This result bears some similarity to the diffusion of happiness—a fuzzy, emotional construct (i.e., risk-tolerance) diffuses through a network. Another researcher has shown that people steal stock ideas from their geographic neighbors (Ivkovic & Weisbenner 2007). Taken together, we see a very interesting mix of things diffusing across networks, from tangible products (Leskovec et al. 2007), to innovative processes (Rogers 1995), to fuzzy concepts like happiness (Fowler & Christakis 2008) and risk-tolerance (Scherer & Cho 2003). This is clearly not all the diffusion literature, which is simply too broad to cover completely here. However, the treatment here covers the classics, while at the same time sampling the state of the art in diffusion work. Next, I examine tie strength's role in diffusion.

### Diffusion and tie strength

Tie strength is almost completely absent from diffusion studies. For reasons of tractability, or perhaps because of lack of data and solid constructs, almost all diffusion studies assign a simple probability distribution to govern when and where information spreads (Fowler & Christakis 2008; Gruhl et al. 2004; Kempe et al. 2003; Kossinets et al. 2008; Liben-Nowell & Kleinberg 2008). (This includes, and perhaps started with, the popular SIR model.) Only one paper breaks the rule. (Another paper seems to make a link between tie strength and diffusion, but actually operationalizes weights on the graph *as* diffusion (Kossinets et al. 2008). They do not start *a priori* from tie strength.) The only diffusion work I am aware of to consider tie strength is the Onnela et al. (2006) study of mobile phone customers. In this large-dataset study, the authors obtained a record of mobile phone calls and induced a network from it (i.e., a phone call creates a link between two people). They operationalize tie strength as minutes of call time. This seems natural and intuitive in this context. As it costs time, and crucially money, to call someone from a mobile phone, call minutes seem like a reasonable proxy for tie strength here. But again, we see a simple proxy substituting for tie strength.

Late in the paper, Onnela et al. *simulates* information flows through a network whose edges understand tie strength. Figure 10 is a reproduction of their simulation. The authors find that global diffusion patterns look very different when weighted by a variable drawn from the "call minutes tie strength distribution." This is valuable work, and it is a perfect foil for the work I present in Chapter 5. A more recent paper from the 2010 WWW conference (De Choudhury et al. 2010) demonstrates the radically different network topologies which arise from different definitions of a tie (e.g., "at least one message between the pair," "at least 10 messages between the pair," or "reciprocal attention by each person"). Yet, still we seem to be somewhat far off from Strang and Soule's call for "closer inspection of the content of social relations between collective actors" in future diffusion research.



Figure 10. A reproduction of a simulated diffusion process based on tie strength, from (Onnela et al. 2006). The simulations show that if tie strength operates on the diffusion process (i.e., diffusion is a function of tie strength), we should expect very different macroscopic properties.

# This work

This dissertation fills the gap. It connects tie strength to actual acts of diffusion on the web. Specifically, I study how tie strengths inferred by We Meddle relate to acts of content forwarding on Twitter via the retweet mechanism. Chapter 5 presents these results. It is a big step for the literature on diffusion because, until now, we have not known how tie strength actually operates on diffusion. I present some surprising findings, such as how tie strength interacts with types of media: text diffuses at a different tie strength rate than images do, for instance. I also examine the content of the forwarded messages and study how tie strength interacts with topics of content pushed around the web.

# CHAPTER 3 COMPUTING TIE STRENGTH

When two people interact via a social medium, like email or Facebook, do they leave recognizable signatures of their closeness? For example, do strong ties communicate more often, in a particular direction or use certain hallmark words? Do weak ties have another signature? If so, can it be recognized with minimal computational cost? Or perhaps the signatures don't exist at all.

In this chapter I describe a study to answer these questions. The study is set in Facebook<sup>5</sup>, a popular social media site with more than 500 million active users<sup>6</sup>. At the outset, it was not clear if tie strength could be reconstructed at all. Indeed, an entirely likely negative hypothesis loomed large: real relationships are too multidimensional and context-dependent to model. This seemed like a entirely valid counter-proposal.

At a high level, the goal of the study presented in this chapter is to map the traces we leave in social media to something both relationally meaningful and important: tie strength. Knowing how to construct tie strength would be an important discovery for theory and for practical applications. I mean "theory" in a somewhat loose sense here: both the theory of tie strength and the theory of tie strength in mediated channels. The hope for practical applications is that the tie strength found in Facebook reflects the tie strength found in social media generally (e.g., email, instant messages, bulletin boards, IRC, USENET, etc.)

Why Facebook? At the time of the study presented here (and at the time of this writing), Facebook played a pivotal role in social media. It is one of the five most visited sites on the internet, with some firms reporting it as the most visited site overall. It also appropriates various features from earlier social media. It has an internal private messaging system that looks like email in many ways. It affords public and semi-public (network-restricted) social communities, in the form of Walls and topical groups, for instance. It has a built-in chat client. The network looms large in Facebook, occupying a pivotal place in how people structure ties and interact online. For these reasons, it seemed reasonable that by studying Facebook I might learn something about these other media from which Facebook draws so heavily. (Of course, that remains to be proved; the next chapter examines it closely.) In these ways, Facebook presented an ideal place to study the properties of mediated tie strength.

This study lays the foundation for how the rest of the dissertation sees and treats tie strength. The results of the study confirm that tie strength can be reconstructed from digital traces, and with surprising

<sup>&</sup>lt;sup>5</sup> http://facebook.com

<sup>&</sup>lt;sup>6</sup>http://www.facebook.com/press/info.php?statistics

accuracy. The *Adjusted-R*<sup>2</sup> value for the model presented here is above 0.5, and on a binary "strong vs. weak" classification task, the model performs with roughly 89% accuracy. The predictive power comes not from the model, which purposely embraces simplicity, but from carefully-drawn, theoretically-meaningful features informed by the tie strength literature. This chapter presents the study, the statistical analysis of its data, and what this model means for both theory and practical applications.

#### **RESEARCH QUESTIONS**

Formally, this chapter addresses two research questions. They are informed by previous scholarship on tie strength in immediate (i.e., face-to-face) contexts.

R1: The existing literature suggests seven categories of tie strength: Intensity, Intimacy, Duration, Reciprocal Services, Structural, Emotional Support and Social Distance. (See Chapter 2 for more.) As manifested in social media, can these categories predict tie strength? In what combination?

R2: What are the limitations of a tie strength model based solely on social media?

R1 embraces the history of the face-to-face tie strength scholarship. The previous chapter covered this literature in detail. For instance, Nan Lin is primarily associated with *Social Distance* tie strength arguments. I answer R1 quantitatively and computationally by analyzing the traces people leave in the social media site Facebook. R2 dissects the tie strength model's mistakes. When does it break down, and why? I study this with post-study interviews.

# THE LANGUAGE AND ARCHITECTURE OF FACEBOOK

Before we move into method, it pays to get some definitions and concepts clear. I will not dwell here, as other articles have detailed the features of social network sites, but since the rest of this chapter adopts some language unique to Facebook and modern social media, I will review them here.

Facebook calls itself a "social utility." It allows people on the internet to friend one another, forming *articulated*, ego-centric social networks. Users extend offers of friendship to one another; if the receiving persons accepts, they become friends. Although you can set privacy parameters for each friendship individually, usually friendship means that the dyad can communicate freely and see most activities the friend performs on the site.

In 2008<sup>7</sup>, Facebook users could write status updates, leave messages on each other's "walls", privately message, post photos and comment on links and photos. Most of this has not changed. Status updates are short text messages that users write about anything and everything, often having to do with ordinary life. Writing on someone's wall amounted to network-visible communication, as Facebook often highlighted these acts to friends in the *Newsfeed*. The Newsfeed, with an interesting history of its own,

<sup>&</sup>lt;sup>7</sup> Facebook has since changed since 2008—more on this later in the chapter.

was (and is) the main way users interact with the site: it is the design around which the site is built. Updates from friends in your network stream to you through the Newsfeed. If I leave a message on your wall, Facebook will permit all of your friends to see it and, most likely, broadcast the message to our mutual friends via the Newsfeed.

# THE SPECIFIC GOAL OF THIS STUDY

This chapter aims to answer the research questions above, but specifically I want to turn the traces we leave in social media into quantified tie strength. In the end, the traces in this study all map onto a raw, real number.

# What does one number mean?

Mapping lots of data into one real number is inherently lossy. We can see this both positive and negative lights. On the positive side: we simplify the world of traces into a realistic and usable quantity from which systems can reason. Intuitively, a single number also generalizes well; we might consider applying it directly in any number of technical situations, like movie recommendations, trust inference, etc. However, it may also over-simplify, making specific situations like movie recommendations too general. Perhaps when a system makes recommendations to users based on friends' preferences, some other slightly different combination of predictors makes more sense. I acknowledge this tension.

This work is a first step. It's an important step because it moves us beyond using simple heuristics, while at the same time providing a general framework from which social systems can reason. As I show in the next chapter, this collapsed tie strength value generalizes to new domains. The idea is that developers can pick up this representation in a new domain in which we cannot immediately reason about tie strength. However, I look forward to new work showing where this model breaks down: contexts where different mixtures make more sense.

ig them

facebook John Doe Fr	iends Applications Inbox Home 🕈 Settings	How strong is your relationship with this person?
	John Doe	barely know them we are very close
	Wall Info Photos	How would you feel acking this friend to lean you \$100
	How strong is your relationship with this person? barely know them we are very close	or more?
	How would you feel asking this friend to loan you \$100 or more?	would never ask very comfortable
View Photos of John (107)	would never ask very comfortable	How helpful would this person be if you were looking
Send John a Message	How helpful would this person be if you were looking for a job?	for a job?
Networks:	no help at all very helpful	no help at all
CUNY Hunter Grad Student '09 Ullinois Alum New York, NY Relationship Status: Married to	How upset would you be if this person unfriended you? not upset at all very upset	How upset would you be if this person unfriended you?
Jane Doe Birthday:	If you left Facebook for another social site, now important would it be to bring this friend along?	Not upset at an very upset
May 19	would not matter must bring them!	If you left Facebook for another social site, how impor-
Current City: Brooklyn, NY	🐺 Write 🔞 Post Photo 🕅 Record Video 🕤 Share Link 🖀 Give Gift	tant would it be to bring this friend along?
Mutual Friends	Write something	would not matter must bring them

Figure 11. The questions used to assess tie strength, embedded into a friend's profile as participants experienced them. An automated script guided participants through a random subset of their Facebook friends. As participants answered each question by dragging a slider, the script collected data describing the friendship. The questions reflect a diversity of views on tie strength.

#### METHOD

Working in our lab in May 2008, I used the Firefox extension Greasemonkey<sup>8</sup> to guide participants through a randomly selected subset of their Facebook friends. (Randomly sampling participants' friends guards against those with large networks dominating the results.) That is, an instrumented laboratory computer collected relational data from each participant. The Greasemonkey script injected five tie strength questions into each friend's profile after the page loaded in the browser. Figure 11 shows how a profile appeared to a participant during a session. Participants answered the questions for as many friends as possible during one 30-minute session. On average, the 35 participants rated 62.4 friends ( $\sigma = 16.2$ ), resulting in a dataset of 2,184 rated Facebook friendships. It's important to note the page looked entirely like a Facebook page participants would usually see on Facebook, except for the addition of the tie strength questions. In this way, the design provides relevant contextual information about a friend (e.g., photos, status updates, biographical information). Participants were free to browse this information during sessions.

Social media experiments often employ completely automated data collection. In other words, very few researchers in this field bring people into the lab. (Collecting data automatically generally creates *much more* data.) I worked in the lab for two important reasons. First, the Greasemonkey script captured all data at the client side, after a page loaded at the user's request. This allowed me to stay within Facebook's Terms of Service (TOS). (Facebook's TOS does not allow a user to scrape the site.) More importantly, I asked participants to provide sensitive information: their relationship strengths plus personal Facebook data. A lab setting guarded participants' privacy and hopefully increased the accuracy of their responses<sup>9</sup>.

# PARTICIPANTS

The 35 participants, primarily students and staff from the University of Illinois community, came from more than 15 different academic departments, such as Law, Medicine and various Engineering departments. I advertised widely across the university to obtain this sample, such as in cafes, libraries and university-wide mailing lists. (40 people originally participated in the study, but I had to discard five participants' data due to a bug in the data collection code.) The sample consisted of 23 women (66%) and 12 men (34%) ranging between 21 and 41 years old, with a mean and median of 26. The minimum number of Facebook friends was 25; the maximum was 729 (median of 153). In terms of age and number of friends, previous work suggests that these participants fall within the mainstream of Facebook

<sup>8</sup> http://www.greasespot.net

<sup>&</sup>lt;sup>9</sup> In 2008, Facebook's API did not yet exist. Today, you can get some of the data presented here through the API, but not all of it. I wanted to run an experiment based on theoretically-inspired data, rather than simply what I could easily obtain via the API.

users (Golder et al. 2007; Rapleaf 2008). All participants used Facebook regularly and had been members for at least one year.

While I expressly tried to obtain a broad sample across the university (i.e., more than simply CS undergrads), this is still a convenience sample. It seems reasonable to assume that some characteristics of relationships hold steady when generalized to broader populations; however, some characteristics certainly may not. Many interesting questions for future work arise simply by reframing the sample. What constitutes tie strength in populations with higher or lower median ages? (For instance, older users have recently adopted Facebook in droves.) What happens to tie strength, if anything, across biographical breaks, such as the transition from college life to professional life? I am excited to see these questions addressed, but leave them for future work.

### PREDICTIVE VARIABLES

While participants responded to the tie strength questions, the browser-based script automatically collected data about the participant, the friend and their interaction history. The tie strength literature reviewed in the previous section pointed to seven major categories of predictive variables. With these categories as a guide, I identified 74 Facebook variables as potential predictors of tie strength. Table 1 presents 32 of the variables I chose along with their distributions. (The full list appears in Appendix A, including interaction terms not counted among the 74.) In choosing these predictive variables, I tried to take advantage of Facebook's breadth while simultaneously selecting variables that could carry over to other social media. Below, I clarify some variables listed in Table 1 and present those not included in the table. All predictive variables make an appearance either in the text or in Table 1, and in full in Appendix A.

Predictive Intensity Variables	Distribution	Max
Wall words exchanged		9549
Participant-initiated wall posts		55
Friend-initiated wall posts		47
nbox messages exchanged		9
nbox thread depth		31
Participant's status updates		80
Friend's status updates		200
Friend's photo comments		1352
ntimacy Variables		
Participant's number of friends		729
Friend's number of friends		2050
Days since last communication		1115
Wall intimacy words	<u> </u>	148
nbox intimacy words		137
Appearances together in photo		73
Participant's appearances in photo		897
Distance between hometowns (mi)		8182
Friend's relationship status	6% engaged 32% married 30% single 0% in relatioBship	
Duration Variable		
Days since first communication		1328
Reciprocal Services Variables		
inks exchanged by wall post	L	688
Applications in common		18
Structural Variables		
Number of mutual friends		206
Groups in common		12
Norm. TF-IDF of interests and about		73
Emotional Support Variables		
Wall & inbox positive emotion words		197
Nall & inbox negative emotion words	L	51
Social Distance Variables		
Age difference (days)		5995
Number of occupations difference		8
Educational difference (degrees)		3
Overlapping words in <i>religion</i>		2
Political difference (scale)		4

Table 1. Thirty-two of over seventy variables used to predict tie strength, collected for each of the 2,184 friendships in our dataset. The distributions accompanying each variable begin at zero and end at the adjacent maximum. Most variables are not normally distributed. The *Predictive Variables* subsection expands on some of these variables and presents those not included in this table.

### Intensity variables

Each Facebook user has a Wall, a public communication channel often only accessible to a user's friends. (Sometimes, but rarely, users permit Wall access to only a select group of friends using privacy settings.) *Wall words exchanged* refers to the total number of words traded between the participant and the friend via Wall posting. *Inbox messages exchanged* counts the number of appearances by a friend in a participant's Facebook Inbox, a private communication channel. *Inbox thread depth*, on the other hand, captures the number of individual Inbox messages sent between the pair. A helpful analogy for *Inbox thread depth* is the number of messages in a newsgroup thread. This is an important distinction that makes an appearance in the Results section.

# Intimacy variables

To complement these aggregate measures, I used the Linguistic Inquiry and Word Count (LIWC) dictionary to perform content analysis (Pennebaker & Francis 1999). While many social scientists are familiar with LIWC, computer scientists may not be. LIWC is a hand-built affective dictionary. It matches text against lists of word stems assembled into categories, and has been validated in many experimental settings. While its validation appealed to me, I chose LIWC primarily because of its computational simplicity. Due the requirements of my experimental design (set forth in the IRB application), I could not store text. So content analysis needed to happen in the browser, in Javascript, without slowing the user's experience. LIWC fit these requirements, while more expressive and computationally intensive approaches (Esuli & Sebastiani 2007; Pang et al. 2002; Thomas et al. 2006) did not.

My hypothesis in applying automatic content analysis was that friends of different tie strengths would use different types of words when communicating. *Wall intimacy words* refers to the number of Wall words matching at least one of eleven LIWC categories: Family, Friends, Home, Sexual, Swears, Work, Leisure, Money, Body, Religion and Health. (I expected the Work category to act as a negative predictor.) Similarly, *Inbox intimacy words* refers to the number of Inbox words matching at least one of these categories. The Home category, for example, includes words like "backyard" and "roommate," while the Work category includes "busy," "classes" and "commute." In total, the intimacy variables checked for matches against 1,635 word stems. Although not presented in Table 1, the model includes each LIWC intimacy category as its own predictive variable.

*Days since last communication* measures the recency of written communication in some Facebook channel (Wall, Inbox, photo comments) from the day I collected data (in May 2008).

#### Duration variable

I did not have access to the date when two people became friends. Instead, *Days since first communication* is a proxy for the length of the friendship. It measures time in the same way as *Days since last communication*. I would have preferred the timestamp of friendship creation. But no one outside Facebook has access to this variable. One of the disadvantages in proxying with *Days since first communication* is the lack of expressivity when two people have only communicated once. In this case, for example one Wall message between the pair, *Days since first communication* equals *Days since last communication*. This is an experimental compromise, but one that I revisit in the next section. Luckily, I was able to obtain a more robust measure of relationship duration in Twitter.

#### Reciprocal services variables

Mark Granovetter identified *Reciprocal Services* as predictor of tie strength. However, he does not go on to define it; later authors have adopted *informational, social* or *economic goods* as its operationalization. In contrast to real life, Facebook friends have relatively few opportunities to exchange these informational, social or economic goods. (These practices clearly differ by media; consider a LinkedIn user who exploits his social capital by introducing business contacts to one another.) To capture *Reciprocal Services* on Facebook, *Links exchanged by wall post* measures the number of URLs passed between friends via the Wall (an information good), a common Facebook practice. Similarly, *Applications in common* refers to the number of Facebook applications a participant and friend share. Facebook applications usually provide a tightly scoped service (e.g., displaying a virtual bookshelf on a profile) and often spread between friends by word of mouth. (Applications appeared on the Facebook site only a few months before my lab sessions. Attitudes and diffusion patterns have clearly changed since then, although I leave this exploration to future work.)

#### Structural variables

Facebook allows users to join groups organized around specific topics and interests. For example, fans of Michael Jackson or people who home-brew beer can join together in a group. The group serves to enable topical conversations, but also functions as an identity marker. *Groups in common* refers to the number of Facebook groups to which both the participant and the friend belong. *Normalized TF-IDF of interests and about* measures the similarity between the free text interests and about profile fields. It does so by computing the dot product between the TF-IDF vectors (the cosine similarity) representing the text. TF-IDF is a standard information retrieval technique (Frakes & Baeza-Yates 1992) that respects the baseline frequencies of different words in the English language. I also measured *Number of overlapping networks*, the number of Facebook networks to which both the participant and the friend belong. Facebook networks often map to universities, companies and geographic areas. Again, because of computational constraints imposed by the experimental design (and the difficulty of obtaining a large corpus of Facebook text), I could not apply more expressive techniques such as Latent Semantic Analysis (Dumais et al. 1988). I revisit the potential for computationally sophisticated techniques like LSA in the upcoming Results section.

### Emotional support variables

In a way similar to the content analysis variables described above, *Wall & inbox positive emotion words* is two variables referring to matches against the LIWC category Positive Emotion. The Positive Emotion category includes words like "birthday," "congrats" and "sweetheart." Similarly, *Wall & inbox negative emotion words* is two variables counting matches in the Negative Emotion category, including words like "dump," "hate" and "useless." I also recorded the number of gifts given between a participant and a friend. A Facebook gift is a small icon often given to a friend to show support. Gifts sometimes cost a small amount of money.

## Social distance variables

I measured the difference in formal education between a participant and a friend in terms of academic degrees. It is computed by searching for the letters BS, BA, BFA, MS, MA, MFA, JD, MD and PhD in the education profile field. *Educational difference* measures the numeric difference between a participant and a friend along a scale: 0: None, 1: BS/BA/BFA, 2: MS/MA/MFA, 3: JD/MD/PhD.

1,261 people in the dataset completed the politics profile field. Of those, 79% reported their political affiliation as "very conservative," "conservative," "moderate," "liberal" or "very liberal." Assigning a scale in that order, *Political difference* measures the numeric difference between a participant and a friend. While the education and politics scales do not completely reflect the diversity of the sample, they do provide useful tools for assessing the importance of these variables for the majority of it.

## Demographic and usage variables

Finally, in addition to the variables described above, I collected demographic and usage information on our participants and their friends: *Gender*, *Number of applications installed*, *Number of inbox messages*, *Number of wall posts* and *Number of photo comments*.

#### DEPENDENT VARIABLES

Previous literature has proposed various manifestations of tie strength (Granovetter 1973; Granovetter 1974; Haythornthwaite 2002; Krackhardt & Stern 1988). To capture a diversity of views, I asked participants to answer five tie strength questions. Participants moved a slider along a continuum to rate a friend. Figure 11 shows how those questions were embedded into a friend's profile. Table 2 illustrates the responses. I chose a continuum instead of a discrete scale for three reasons. First, Granovetter conjectured that tie strength may in fact be continuous. The literature has not resolved the issue, let alone specified how many discrete tie strength levels exist. A continuum bypasses that problem. Second, a continuum lends itself to standard modeling techniques. Finally, applications can round a continuous model's predictions to discrete levels as appropriate.

Often, social psychologists will combine participant responses into a scale using a weighted linear combination. It is a also common to use tested instruments that conform to internal and external valid-

ity measures. However, this setting (i.e., socializing via the internet) breaks the assumptions of the few existing instruments. It also seemed inappropriate to build a scale because, although previous face-to-face literature expresses tie strength in these terms (e.g., "How helpful when looking for a job?"), it remains completely unclear *a priori* whether this means the same thing in mediated environments. Therefore, it seemed prudent to apply disaggregation, while at the same time leaning on the literature.

Dependent Variables	0–1 Scale	Median
How strong is your relationship?	li.ii	0.411
How comfortable asking for loan?		0.076
How helpful if looking for job?	<b> </b>	0.362
How upset if unfriended?	<b>            </b>	0.552
How important to bring friend?	<b> </b>	0.324

Table 2. The five questions used to assess tie strength, accompanied by their distributions. The distributions present participant responses mapped onto a continuous 0–1 scale. This model predicts these responses as a function of the variables presented in Table 1.

#### STATISTICAL METHODS

I model tie strength as a linear combination of the predictive variables, plus terms for category interactions and network structure:

$$s_{i} = \alpha + \beta R_{i} + \gamma D_{i} + N(i) + \epsilon_{i}$$

$$N(i) = \lambda_{0}\mu_{M} + \lambda_{1}med_{M} + \sum_{k=2}^{4} \sum_{s \in M} \lambda_{k}(s - \mu_{M})^{k}$$

$$+ \lambda_{5}min_{M} + \lambda_{6}max_{M}$$

$$M = \{s_{i} : j \text{ and } i \text{ are mutual friends}\}$$

In the equations above,  $s_i$  represents the tie strength of the  $i^{th}$  friend.  $R_i$  stands for the vector of 74 individual predictive variables.  $\varepsilon_i$  is the error term.  $D_i$  represents the pairwise interactions between the categories presented in Table 1. Pairwise interactions are commonly included in predictive models (Gergle et al. 2006); in this case, including all pairwise interactions would force more variables than data points into the model. Instead, I nominated variables with the fewest missing values to represent each category. (Not every participant or friend contributes every variable.)  $D_i$  represents all pairwise interactions between the 13 variables with a 90% or greater completion rate. Choosing 90% as a threshold ensured that every category was represented. To the best of my knowledge, existing work has not explored interactions between the categories of tie strength. This model does not include fixed effects—effects specific to individuals. Some communities, for example Economics, often include fixed effects in models like this one. It is possible that fixed effects could offer explanatory power beyond this random effects model. I opted for a random effects to offer explanatory power without the need to fallback on individual-level explanations, which would severely limit a tie strength model's applicability in other domains.

More complex models were explored, but a (mostly) linear model takes advantage of the full dataset and can explain the results once it is built. Whereas train-test methods have purchase in many machine learning communities, the social sciences often opt for constrained models (e.g, the mostly linear one above) that take advantage of entire datasets and offer explanations for theory. For instance, I explored SVM regression (Flake & Lawrence 2002) to model tie strength. The fit was fantastic: the *R*<sup>2</sup> was above 0.9. In computer science venues, I would typically report this number. (As I show momentarily, the fit for the model presented here is considerably worse than this SVM fit.) However, a complex technique like SVM regression has two major drawbacks. First, it does not allow us to actually see what makes tie strength. From our point of view, it is a black box. Second, in the next chapter I present an application of the model in a new domain. I had to weigh goodness of fit against likelihood of generalization. In this case, I opted for simplicity (e.g., the linear model) over a very good, but complex fit.

# Network effects

N(i) encodes network structure. It captures the idea that a friendship's tie strength not only depends on its history, but also on the tie strengths of mutual friends. In other words, it models the idea that a friend who associates with your business acquaintances is different than one who knows your mother, brother and sister. Since every friend has a potentially unique set of mutual friends, the model uses seven descriptors of the tie strength distribution over mutual friends: mean, median, variance, skew, kurtosis, minimum and maximum. These terms belong to the *Structural* category. However, N(i) introduces a dependency: every tie strength now depends on other tie strengths. It's now recursive.

How can the model incorporate the tie strengths of mutual friends when it is tie strength it wants to model in the first place? To solve this problem, I fit the equations above using iterative (OLS) regression. In each iteration, the tie strengths from the previous round are substituted to calculate N(i), with all  $s_i$  initially set to zero. Using this procedure, all  $s_i$  converged in nine iterations (0.001 relative change threshold for every  $s_i$ ). This approach parallels other "neighborhood effect" models (Chopra et al. 2007). A common alternative modeling technique is a network regularization framework, in which the network effects do not play a role in minimizing error during the model fit stage. Instead, you first express tie strength as best you can strictly from interaction data, and then follow it with a step that smooths all  $s_i$  in the same clique. From a purely computational perspective, this approach is attractive. However, it disconnects model-fitting from participant responses. What function should smooth  $s_i$  in the same clique? If it's learned, it appears roughly equivalent to the procedure described above. Furthermore, any function for smoothing cliques runs the risk of overfitting. (There is a strong theoretical

case here for limiting features to a small basket of meaningful statistical summaries.) Finally, a regularization framework would have difficulty interacting with raw relationship data. While certainly not impossible,  $D_i$  would be difficult to express in this framework. (Additionally, a model that first decomposes the network into communities before applying N(i) seems like an attractive route for future modeling work. I revisit this idea in *We Meddle*.)

# Standardization and normalization

I did not standardize, or "ipsatize" (Cunningham et al. 1977), the dependent variables. Because I employed network subsampling, I could not be sure participants saw the Facebook friend they would rate highest or lowest. Ipsatizing the dependent variables in this way may have made them easier to model; however, it would also stretch them in a way almost certainly at odds with participants' judgments. Furthermore, not all real-life friends have Facebook accounts. It is reasonable to assume that some participants would reserve the ends of the spectra for people our experiment would never uncover, due to the randomized experimental procedure.

Finally, to account for the violations of normality exhibited by the distributions in Table 1, every variable is log-transformed. (This does not fully control the violations of normality. The Results section applies robust standard errors to control for it.)



Figure 12. The model's Adjusted R<sup>2</sup> values for all five dependent variables, broken down by the model's three main terms (added in sequence from top). Modeling interactions between tie strength categories results in a substantial performance boost. The model performs best on *Loan* \$100? and *How strong*?, the most general question.

# RESULTS

Because each participant rated more than one friend, observations within a participant were not independent. This is a common obstacle for ego-centric designs. To roughly adjust for it, all of the results presented here cut the degrees of freedom in half, a technique borrowed from the social networks literature (Marsden & Campbell 1984). This ensures greater skepticism in the face of a random, rather than fixed effects model.

On the first tie strength question, *How strong is your relationship with this person*?, the model fits the data very well: *Adj.*  $R^2 = 0.534$ , p < 0.001. It achieves a Mean Absolute Error of 0.0994 on a continuous 0–1 scale, where 0 is weakest and 1 is strongest. In other words, on average the model predicts tie strength within one-tenth of its true value. This error interval tightens near the ends of the continuum because predictions are capped between 0 and 1. In addition, I found strong evidence of four category interactions (p < 0.001): *Intimacy* × *Structural*,  $F_{1,971} = 12.37$ ; *Social Distance* × *Structural*,  $F_{1,971} = 34$ ; *Reciprocal Services* × *Reciprocal Services*,  $F_{1,971} = 14.4$ ; *Structural* × *Structural*,  $F_{1,971} = 12.41$ . As I dem-

onstrate shortly, the *Structural* category plays a minor role as a linear factor. However, it has an important modulating role via these interactions. One way to read this result is that individual relationships matter, but they get filtered through a friend's clique before impacting tie strength. (I use the word "clique" loosely here, rather than in its more precise, graph-theoretic meaning.)

Figure 12 summarizes the model's performance on all five tie strength questions, broken down by the model's three main terms. Modeling category interactions boosts performance significantly, with smaller gains associated with modeling network structure. (The performance boosts are accounted in that order, with interactions preceding network structure. In other words, the boost associated with structure accounts for that gains via interactions, but not vice versa.) The model fits the second tie strength question as well as the first: How would you feel asking this friend to loan you \$100 or more? However, it does not fit the last three questions as well. The lower performance on these questions may have resulted from participant fatigue. (Participants, after all, moved approximately 312 sliders during a typical session.) I considered randomizing the questions for each friend to account for ordering effects like fatigue (as many experimental designs would), but I feared that randomizing would confuse and frustrate participants, contributing to lower accuracy across the board. Therefore, I chose to prioritize the first question, the most general of the five. (This is also the model I take along to the next section.) With the exception of How helpful would this person be if you were looking for a job?, all dependent variable intercorrelations were above 0.5 (Table 3). The high intercorrelations indicate that, on the whole, the dependent variables measure related concepts. However, as I demonstrate shortly, the predictors useful in one context do not necessarily help predict another.



Figure 13. The predictive power of the seven tie strength categories, presented here as part of the *How strong?* model. A category's weight is computed by summing the absolute values of the coefficients belonging to it. The diagram also lists the top three predictive variables for each category. On average, the model predicts tie strength within one-tenth of its true value on a continuous 0–1 scale.

Figure 13 visualizes the predictive power of the seven tie strength categories as part of the *How strong*? model. The figure also includes each category's top three contributing variables. The weight of a category is calculated by summing the coefficients of the the variables belonging to it. Although not uniformly distributed, no one category has a monopoly on tie strength. Figure 13 first appeared in (Gilbert & Karahalios 2009) and it does not adequately deal with multicollinearity. For instance, in certain cases, it takes the absolute values of (negatively) correlated variables and adds them together. In the Discussion section, I discuss the implications of a more sophisticated Principal Components Analysis. PCA uncovers four roughly uncorrelated "true" dimensions of tie strength, as opposed to the seven outlined above. At the same time, this may be unfair. For example, *Emotional Support* is particularly hard to capture. I relied on a linguistic analysis of users' posts, which correlates highly with other, coarser measures. It seems premature to conclude that *Emotional Support* does not exist, given the difficulty in capturing it fully in social media. I comment further on these effects in the upcoming Discussion section.



Figure 14. The model's performance across all ties in the dataset. There is a strong correlation, with Adj. R<sup>2</sup> above 0.5. Yet the model shows a very slight bias toward underestimation, represented as the larger cloud in the bottom-right of the figure. The gap in the center results from participants' inclination to move the slider from its starting point, if only slightly.

Table 4 presents the standardized beta coefficients of the top predictive variables on the *How strong?* question. The *F*-statistics signify a variable's importance in the presence of the other variables (using R's leave-one-out ANOVA procedure, called drop1). These *F*-statistics control for multicollinearity by testing the variance explained by a model missing a predictor against the full model. The resulting *F*-statistics show the power, or impact, of a particular predictive variable. The two *Days since* variables have such high coefficients due to friends that never communicated via Facebook. (They are also highly correlated, as cases of 0 or 1 interaction result in the same *Days since* number.) Those observations were assigned outlying values: zero in one case and twice the maximum in the other. In other words, the simple act of communicating once leads to a very large movement in tie strength. *Educational difference* plays a large role in determining tie strength, but that may reflect the university community from which I sampled participants. Curiously, *Inbox thread depth* has a negative effect on tie strength; the more messages friends exchange on a single topic, the lower their tie strength. It is important to note that Table 4 orders the variables by their weights, or  $\beta$  coefficient; it expresses confidence in its utility relative to other variables (as measured by leave-one-out ANOVA). (The coefficient confidence is

greater than 99.9%.) For example, *Inbox thread depth* is highly correlated with *Inbox intimacy words*, resulting in a lower *F*-statistic.

Figure 14 compares the model's prediction to participant responses across the entire dataset. The figure illustrates a strong correlation and another view on the MAE presented above. Figure 15, a complement to the preceding figure, recasts the original regression problem as a two-class classification problem. It divides the scatterplot into quadrants, with lower-left and upper-right quadrants representing correct predictions of strong tie and weak tie, respectively. The other quadrants represent errors. The *How strong*? model classifies with 88.7% accuracy using this procedure, significantly outperforming the baseline,  $\chi^2(1, N = 4368) = 700.9$ , p < 0.001.



Figure 15. The data from the previous figure, recast as a binary classification problem. *Weak tie* is interpreted as those ties participants rank as below average, and *Strong tie* as those above the mean. (x-axis  $\mu$ =0.43 and y-axis  $\mu$ =0.46.) Under this simple reformulation, the model performs at roughly 89% classification accuracy.

Correlations	Strong	Loan	Job	Un	Bring
Strong	1	0.69	0.45	0.75	0.7
Loan	0.69	1	0.4	0.55	0.55
Job	0.45	0.4	1	0.5	0.46
Unfriend	0.75	0.55	0.5	1	0.74
Bring	0.7	0.55	0.46	0.74	1

Table 3. The intercorrelations of the five dependent variables. With the exception of Job-
Strong, Job-Loan and Bring-Job, the dependent variables are well-correlated with one another

Top Predictive Variables (How strong?)	β	F	p-value
Days since last communication	-0.762	453	< 0.001
Days since first communication	0.755	7.55	< 0.001
Intimacy × Structural	0.4	12.37	< 0.001
Wall words exchanged	0.299	11.51	< 0.001
Mean strength of mutual friends	0.257	188.2	< 0.001
Educational difference	-0.223	29.72	< 0.001
Structural × Structural	0.195	12.41	< 0.001
Reciprocal Serv. $ imes$ Reciprocal Serv.	-0.19	14.4	< 0.001
Participant-initiated wall posts	0.146	119.7	< 0.001
Inbox thread depth	-0.137	1.09	0.29
Participant's number of friends	-0.136	30.34	< 0.001
Inbox positive emotion words	0.135	3.64	0.05
Social Distance × Structural	0.13	34	< 0.001
Participant's number of apps	-0.122	2.32	0.12
Wall intimacy words	0.111	18.15	< 0.001

Table 4. The top predictive variables as measured by standardized beta coefficients. The two *Days since* variables have large coefficients because of the difference between never communicating and communicating once. The utility distribution of the predictive variables forms a power-law distribution: with only these fifteen variables, the model has over half of the information it needs to predict tie strength. *F*-statistics reflect the impact of a particular feature, in this table and in each upcoming predictive variable table.

Top Predictive Variables (Job?)	β	F	p-value
Days since first communication	0.871	4.14	0.042
Days since last communication	-0.843	3.97	0.046
Educational difference	-0.393	19.5	< 0.001
Participant's number of friends	0.310	24.96	< 0.001
Wall words exchanged	0.281	4.59	0.032
Occupational difference (number)	-0.232	4.62	0.032
Participant's inbox messages	0.174	23.27	< 0.001

Wall intimacy words	-0.159	1.65	0.2
Participant's links broadcast	0.138	13.62	< 0.001
Median strength of mutual friends	0.132	1.54	0.22
Participant's number of apps	-0.123	0.06	0.81

Table 5. The top predictive variables as measured by standardized beta coefficients for *How helpful would this person be if you were looking for a job?* Light gray rows indicate variables of particular interest.

Top Predictive Variables (Bring friend?)	β	F	p-value
Educational difference	-0.599	54.11	< 0.001
Occupational difference (number)	0.369	13.93	< 0.001
Participant's number of apps	0.348	39.39	< 0.001
Wall words exchanged	0.340	7.98	0.005
Days since first communication	0.278	0.502	0.48
Days since last communication	-0.250	0.417	0.522
Friend is engaged or married	-0.24	16.97	< 0.001
Participant's number of status updates	-0.238	27.92	< 0.001
Mean strength of mutual friends	-0.205	0.594	0.44
Participant's inbox messages	0.180	29.70	< 0.001
Inbox positive emotion words	0.156	3.87	0.049

Table 6. The top predictive variables as measured by standardized beta coefficients for *If you left Facebook for another social site, how important would it be to bring this person along?* Light gray rows indicate variables of particular interest.

Top Predictive Variables (Loan \$100?)	β	F	p-value
Days since last communication	-0.651	2.891	0.089
Days since first communication	0.651	2.817	0.093
Mean strength of mutual friends	0.304	1.178	0.28
Wall words exchanged	0.299	6.31	0.012
Occupational difference (number)	0.275	7.91	0.005
Friend is male	0.247	0.417	0.522
Friend is engaged or married	0.218	55.84	< 0.001
Educational difference	-0.186	5.33	0.021
Number of mutual friends	0.181	30.84	< 0.001
Participant's number of status updates	-0.159	12.59	< 0.001
Minimum strength of mutual friends	-0.150	1.18	0.28

Table 7. The top predictive variables as measured by standardized beta coefficients for *How would you feel asking this friend to loan you \$100 or more?* Light gray rows indicate variables of particular interest.

Educational difference	-0.797	87.96	< 0.001
Participant's number of friends	0.679	134.08	< 0.001
Days since first communication	0.606	2.19	0.14
Days since last communication	0.604	2.23	0.14
Maximum strength of mutual friends	0.461	11.50	< 0.001
Mean strength of mutual friends	-0.325	1.20	0.27
Wall words	0.321	6.52	0.012
Participant's number of wall posts	-0.298	51.06	< 0.001
Std. dev. of strength of mutual friends	-0.289	13.97	< 0.001
Occupational difference (number)	-0.244	5.57	0.018
Kurtosis of strength of mutual friends	-0.202	12.98	< 0.001

Table 8. The top predictive variables as measured by standardized beta coefficients for *How upset would you be if this person unfriended you?* Light gray rows indicate variables of particular interest.

# WHEN THE MODEL BREAKS DOWN: HIGH RESIDUALS

The model performs well, but not perfectly. When does it break down, and why? Are there structural markers for error? To understand its limitations, I conducted ten follow-up semi-structured interviews about the friendships the model had the most difficulty predicting. After identifying the friends with the highest residuals (for the *How strong*? model), I asked participants to tell me about this particular friendship, including anything that makes it special. For instance, one participant described a "friend" he barely knew:

I don't know why he friended me. But I'm easy on Facebook, because I feel like I'm somehow building (at least a minuscule amount of) social capital, even when I don't know the person. I went to the same high school and have a few dozen common friends. We've never interacted with each other on Facebook aside from the friending.

### participant's rating: 0; prediction: 0.44

Notice how the participant recalls that "he friended me." Although these friends had communicated via Facebook only twice (the participant mistakenly recalled "never"), the friend's clique confused the model. The friend came from a group of relatively strong friends. As I mentioned earlier, the model filters individual relationships through cliques, leading to the high residual. Perhaps having deeper network knowledge could help, such as how the mutual friends see this friend. For instance, taking the judgements of others could help predict this individual's tie strength prediction. This is common network analysis, but it's beyond this particular ego-centric experimental design.

#### ASYMMETRIC FRIENDSHIPS

Two participants rated a friend highly because of how the friendship compared to others like it. In one case, a participant described a close bond with a professor:

This is a professor from one of the classes I TA-ed. We have a very good relationship, because in the past we have worked out a lot of difficult class problems. The professor still remembers my name, which for some of my "friends" on Facebook may not be true. But not only that, she also knows how things are going at school, and when we meet in a hallway we usually stop for a little chat, rather then exchanging casual "Hi! Hello!" conversation.

# participant's rating: 0.85; prediction: 0.41

*Educational difference* and the directionality of the wall posts pushed this prediction toward weak tie. Many people would not remark that a close friend "remembers my name." However, in the context of this participant's "networking" friends, the professor breaks the mold.

Participants' responses often revealed the complexity of real-life relationships, both online and offline. One participant grounded her rating not in the present, but in the hope of reigniting a friendship:

Ah yes. This friend is an old ex. We haven't really spoken to each other in about 6 years, but we ended up friending each other on Facebook when I first joined. But he's still important to me. We were best friends for seven years before we dated. So I rated it where I did (I was actually even thinking of rating it higher) because I am optimistically hoping we'll recover some of our "best friend"-ness after a while. Hasn't happened yet, though.

participant's rating: 0.6; prediction: 0.11

#### CONFOUNDING THE MEDIUM

As might be expected, Facebook friends do not only stick to Facebook. (In fact, earlier work has argued that strong ties often span multiple media (Haythornthwaite 2002), although very recent work has questioned this finding (Hogan 2009).) One participant described a close friendship with a diverse digital trail:

This friend is very special. He and I attended the same high school, we interacted a lot over 3 years and we are very very close. We trust each other. My friend are I are still interacting in ways other than Facebook such as IM, emails, phones. Unfortunately, that friend and I rarely interact through Facebook so I guess your predictor doesn't have enough information to be accurate. **participant's rating: 0.96; prediction: 0.47** 

However, even friends that stick to Facebook sometimes do so in unexpected ways:

We were neighbors for a few years. I babysat her child multiple times. She comes over for parties. I'm pissed off at her right now, but it's still 0.8. ;) Her little son, now 3, also has an account on Facebook. We usually communicate with each other on Facebook via her son's account. This is our "1 mutual friend."

#### participant's rating: 0.8; prediction: 0.28

This playful use of Facebook clearly confused the tie strength model. With the exception of the *Social Distance* category, all indicators pointed to a weak tie. In fact, it is hard to imagine a system that could ever (or should ever) pick up on scenarios like this one. It is clear that users sometimes go to great lengths to subvert systems—a natural practice.

On the whole, these two themes *Asymmetric Friendships* and *Confounding the Medium* stood out from the interview data. I have excerpted relevant passages from that data. Deeper network knowledge may help resolve the first; knowledge across various social media (e.g., email, IM, Twitter, etc.) may help to resolve the second (although this approach is fraught with privacy concerns).

#### DISCUSSION

These results show that social media can predict tie strength. The *How strong*? model predicts tie strength within one-tenth of its true value on a continuous 0-1 scale, a resolution probably acceptable for most applications. In other words, discretizing the tie strength continuum onto a 10-point Likert scale, the How strong? model would usually miss by at most one point. The Intimacy category makes the greatest contribution to tie strength, accounting for 32.8% (in raw, disaggregated terms) of the model's predictive capacity. This parallels Marsden's finding that emotional closeness best reflects tie strength (Marsden & Campbell 1984). However, the Intensity category also contributes substantially to the model, contrasting with Marsden's finding that Intensity has significant drawbacks as a predictor. One way to explain this discrepancy is that the sheer number of people available through social media strengthens Intensity as a predictor. In other words, when you choose to interact with someone over and over despite hundreds of people from which to choose, it significantly informs tie strength. The number of variables representing each category also plays a role in its overall impact. For example, Emotional Support might impact tie strength more if more variables represented it. Emotional Support is particularly hard to quantify, and more sophisticated techniques may do a better job of capturing it more robustly. For instance, semi-supervised learning that starts from a labeled, hand-coded corpus and expands to fit Facebook may help. Or, less sophisticated but clever techniques may finally do the trick, similar to what I did with "birthday wishes." However, more variables does not always equal greater impact. As Duration illustrates, a single variable can account for a large part of the model's predictive capacity.

Some applications will not need 10-point resolution; the coarse categories of *strong* and *weak* may suffice. In "The Strength of Weak Ties," Granovetter himself performs his analytic work with only these approximate distinctions. One way to accomplish this is to use the model's mean, classifying all friends above it as *strong* and all below it as *weak*. Correct predictions are those where the participant's rating is correspondingly above or below the mean in the participant dataset. (With the continuous data supplied by our participants, using means to recast the regression problem as a classification problem seemed like conservative approach.) The *How strong*? model classifies with 88.7% accuracy using this procedure, significantly outperforming the baseline,  $\chi^2(1, N = 4368) = 700.9$ , p < 0.001. (Note that this situation does not require more sophisticated evaluation techniques, like cross-validation, because the model is highly constrained and the threshold is not learned.) Is roughly 89% good? Or could heuristics do an equally good job? For instance, even in my own previous work (Gilbert et al. 2008), heuristics have acted as placeholders. (In the case of Gilbert et al. (2008), the heuristic was messages sent to the friend.) Before this study, we actually did not know the answer to this question. Without reliable data on ground truth tie strength, judging the efficacy of heuristics remained little more than guesswork. Using data from this study, I can now judge the effectiveness of various heuristics. Take one very common heuristic: messages sent to a friend (the one I used in earlier work). It performs with 61% accuracy on the coarse classification task, compared to a 52% baseline (i.e. the strategy "choose weak tie," the most common class). Compared with the full model's 89% accuracy, it's clear that modeling various aspects of tie strength with much resolution in previous work.

The error analysis interviews illustrate the inherent complexity of some relationships. They also point the way toward future research. A model may never, and perhaps should never, predict some relationships. Wanting to reconnect with an ex-boyfriend comes to mind. Relationships like these have powerful emotions and histories at play. However, it may be possible to make better predictions about relationships like the professor-student one, a strong relationship relative to others like it. Incorporating organizational hierarchy may also improve a system's ability to reason about relationships like these. Merging deeper network knowledge with data about who extended the friend request also looks promising, as evidenced by the "he friended me" interview.

#### Individual predictors

Some predictive variables surprised me. For instance, *Inbox thread depth* negatively affects tie strength (although it's *F*-statistic shows that it correlates well with other variables, as it perhaps expected). This finding clashes with existing work. Whittaker et al. (1998) report that familiarity between Usenet posters increases thread depth. One way to resolve this disparity is to note that there may be a fundamental difference between the completely private threads found on Facebook (essentially a variant of email) and Usenet's completely public ones. Common ground theory (Clark 1993) would suggest that strong ties can communicate very efficiently because of their shared understanding, perhaps manifesting as shorter Inbox threads. There is, however, another interpretation. Media multiplexity would suggest that strong ties communicate in varied media (Haythornthwaite 2002). Perhaps ties that rely so heavily on Facebook do so at the expense of other media. In plainer terms, if I need the Facebook inbox, perhaps I

cannot contact you via email, indicating a weak tie. Alas, the the tie strength model does not tell us much about causality here. Future work could tease this apart.

*Together in photo, Normalized TF-IDF of interests* and *Normalized TF-IDF of about* did not turn out to be useful predictors. This is somewhat surprising, especially the case of *Together in photo*, which seems like a very good intuitive predictor. *Together in photo* just contained too little information, rendering it an ineffective predictor: less than 10% of data points contained a nonzero *Together in photo* value. In retrospect, the *Normalized TF-IDF* predictors probably suffer from a different problem. Since these predictors measure word-level similarity, instead of concept-level similarity, it underestimates (probably significantly) the degree of correspondence between two friends' free text descriptions of themselves. A more sophisticated technique like Latent Semantic Analysis may correct this. I applied cosine similarity because a standard, if relatively unsophisticated technique, may have worked: look at the word-count intimacy measures. In this case, however, it seems that a more sophisticated technique may be necessary. (I cannot test this hypothesis as the experimental protocol required me to dispose of text after collecting it.)

*Educational difference* also strongly predicts tie strength, with tie strength diminishing as the difference grows. This may have resulted from the university community to which our participants belonged. On the other hand, the result may have something to do with Facebook itself, a community that spread via universities. Some variables I suspected to impact tie strength did not. *Number of overlapping networks* and *Age difference*, while intuitively good predictors, also made little appreciable difference to tie strength. ( $\beta = 0.027$ ,  $F_{1,971} = 3.08$ , p = 0.079 and  $\beta = -0.0034$ ,  $F_{1,971} = 10.50$ , p = 0.0012, respectively.)

### Different kinds of tie strength

Taken together, the predictive variable tables (Tables 5 through 8) illustrate the underlying structural differences between the five dependent variables. You also see commonalities. Recency and duration appear in all models. At least one *Social Distance* variable, one *Intensity* variable and one *Structural* variable show up in each model, as well. It seems that no model is complete without them.

However, the differences between the models may shed the most light on them. Tables 5 through 8 highlight in gray the variables I find particular interesting. For instance, in the *Job?* model, *Occupa-tional difference* and *Educational difference* figure particularly prominently, most likely reflecting the context of the question. (*Job?* is also the most difficult to predict, achieving an *Adj.*  $R^2 = 0.39$ .) The *Bring friend?* model puts *Educational difference* at the top by a wide margin, it's highest place among all the models. The *Bring friend?* model suggests that people more educated than you and people in serious relationships (*Friend is engaged or married*) do not need to come along to a new site. Also, a curiosity pops up in the *Bring friend?* model: the largest recursive, structural term is negative; the *Mean strength of mutual friends* coefficient is -0.205. This is an anomaly among the models. Perhaps an explanation is

the following: people from strong cliques have multiple media and therefore it decreases the relevance of any one medium. I hope to see future work examine this issue more closely.

The Loan \$100? model has two particularly interesting standouts: Friend is male and Friend is engaged or married. Other than these two variables, Loan \$100 looks quite similar to How strong?, with the difference that you look to married men if you need money. In Upset if unfriended?, three new structural variables pop into the top spots, ones we have not seen in any other models: Maximum strength of mutual friends, Standard deviation of strength of mutual friends and Kurtosis of strength of mutual friends. The nonlinearity of these variables suggests some network-based, rather than dyad-based, complexity. Upset if unfriended? is the one explicitly negative dependent variable. It introduces the possibility of a negative link (i.e., cutting the tie), instead of a merely weak one. In this way, the key structural findings, and their nonlinearity, resonate with the scholarship of structural balance (Heider 1946) and signed networks (Leskovec et al. 2010). In this literature, triads and clusters have structural properties that favor certain configurations over others. Some configurations have stability; some quickly erode with time. Upset if unfriended? seems to tell a similar story.

#### CONCLUSION

The major finding of this work is that tie strength can be reconstructed with high accuracy using the digital traces people leave behind in social media. The development of tie strength here lays the foundational for the rest of the dissertation. In Chapter 4, I show how a web application can use tie strength at the heart of its design. In Chapter 5, I show how to apply this model as a critical component in an analysis.

### Theoretical implications

There is still more variance to understand. Certainly, more predictive variables could help, such as "behind-the-scenes" data like who friended who. However, throwing more data at the problem might not solve it; perhaps social media needs novel indicators. This raises new questions for theory. When modeling tie strength exclusively from social media, do we necessarily miss important predictors? What is the upper limit of tie strength predictability?

After the study, I wished I had included two other particular predictive variables: *politeness* and *intermessage response time*. Politeness could be measured linguistically, although it seems that no current affective dictionary provides very good (if any) coverage of politeness. One approach might start from a small number of examples drawn from various affective dictionaries and theory (e.g., proper titles, hedges, "thank you," etc.), and grow that set using co-training. Inter-message response time is more straightforward to measure, but still might provide a lot of bang for the buck. The canned example I like to use is an email from the boss. Many people run to their machines to compose intricate replies to their bosses. The reverse is almost never true, signaling the asymmetric power relationship. I would recommend that any follow-up work include these two variables in a tie strength model.

I believe that this work makes three important contributions to existing theory. First, I defined the importance of the categories of tie strength as manifested in social media. This is novel especially in light of the fact that these weights do not always align with prior work. Second, I showed that tie strength can be modeled as a continuous value. Third, these findings reveal how the *Structural* category modulates other categories by filtering individual relationships through cliques. Previously, it was not well-understood how or if tie strength categories interacted.

Finally, I see a home for our results in social network analysis. Most work to date has assumed a present link or an absent link, omitting properties of the link itself. Introducing a complete tie strength model into social network analyses, perhaps even joining a social media model with real-world data, may enable novel conclusions about whole systems (Laumann et al. 1989).

### **Practical implications**

I also foresee many opportunities to apply tie strength modeling in social media. Consider privacy controls that understand tie strength. When users make privacy choices, a system could make educated guesses about which friends fall into trusted and untrusted categories. This might also depend on media type, with more sensitive media like photos requiring higher tie strengths. The approach would not help users set privacy levels for brand new friends, ones with whom there is no interaction history. Yet, it has two main advantages over the current state of the art: it adapts with time, and it establishes smart defaults for users setting access levels for hundreds of friends.

Or, imagine a system that only wants to update friends with novel information. Broadcasting to weak ties could solve this problem. Consider a politician or company that wants to broadcast a message through the network such that it only passes through trusted friends. Because strongly tied friends of-ten reconcile their interests (Granovetter 1973), a politician might look for new supporters among the strong ties of an existing one. Limiting the message's audience in this way may increase the success rate relative to the effort expended.

Social media has recently started suggesting new friends to users. However, sometimes I choose not to friend someone with good reason. For instance, a strong tie of a strong tie is not necessarily a friend at all: consider the beloved cousin of a best friend. Granovetter writes, "if strong ties A–B and A–C exist, and if B and C are aware of one another, anything short of a positive tie would introduce a 'psychological strain' into the situation" (Granovetter 1973). A system that understands tie strength might avoid "strain" by steering clear of these delicate situations. In fact, weak ties of existing friends may make better friend candidates, as it is less likely that users have already declined to friend them. More broadly,

systems that understand tie strength might apply it to make better friend introductions, although deeper study would need to uncover how best to use it in this context.

Recent work suggests that the average number of social media friends continues to grow, currently above 300 (Lampe et al. 2008). With users keeping so many friends, social media has started to consolidate friend activity into a single stream. Facebook calls this the Newsfeed. However, the multiplicative nature of the types of friends crossed with the types of updates, e.g., photos, status, new friends, comments, etc., presents a difficult design problem. A system that prioritizes via tie strength, or allows users to tune parameters that incorporate tie strength, might provide more useful, timely and enjoyable activity streams. I explore these ideas with my next project, a Twitter client called *We Meddle*.
## CHAPTER 4 DESIGNING AROUND TIE STRENGTH

After I presented the findings in Chapter 3, two crucial questions emerged. First, to what extent does the model in Chapter 3 capture tie strength as expressed *only in Facebook?* In other words, does it generalize? If so, to what? A model that generalizes has much greater theoretical value than one that does not. For instance, if you want to study email, can the tie strength model help you?

But next came a more practical question. Can you do anything with it? How can a model of tie strength help people who build or design social media? Can it change the way we architect social media systems? For instance, you could argue that since we understand tie strength, we could fundamentally revolve systems around it. Imagine a social media site that ordered status updates not by time, but instead bins them by tie strength: strong ties here, medium ties there and weak ties initially hidden. How would users react? This work began from a theoretical problem: Why don't we understand relationships? However, now that we understand tie strength, substantial social media design challenges now seem tractable.

In this chapter I present a social media site called *We Meddle* that aims to answer both the theoretical and practical question simultaneously. In other words, We Meddle wants to answer the generalizability question, but at the same time it simply wants to make social media a little bit better. We Meddle is a website designed for Twitter<sup>10</sup> users and has been available since early 2010 at http://wemeddle.com. At the time of this writing, over 1,300 people from more than 20 countries have logged in; it has close to 10,000 page views. Users receive no compensation. They log in simply because they want to use We Meddle. Its central feature is that it automatically infers tie strength between you and the Twitter users you by applying the model presented in Chapter 3. We Meddle takes on the *collapsed context problem*, a problem I describe in detail shortly. In this chapter, I present the problem We Meddle aims to solve, as well as its design, deployment and evaluation. The results show that computational tie strength generalizes to a community beyond its initial testbed, and that We Meddle solves problems some users have with social media.

#### **RESEARCH QUESTIONS**

Formally, this chapter addresses three primary research questions.

**R1**: Does the computational tie strength model presented in Chapter 3 generalize to another social site? If so, under what assumptions? If not, where does the model break down?

<sup>&</sup>lt;sup>10</sup> http://twitter.com is a very popular microblogging (i.e., status messaging) site.

- **R2**: Can a computational tie strength model help solve design challenges introduced by the collapsed context problem?
- **R3:** Can feedback from users improve the model? If so, how does it change? (This is known in the machine learning literature as interactive training.)

#### THE LANGUAGE AND ARCHITECTURE OF TWITTER

Whereas in the previous chapter I worked in Facebook, here the context is Twitter. There are two reasons for the switch. First, Twitter makes it much easier to build applications for its service than Facebook. It provides clean, well-documented APIs and (unlike Facebook) takes a very open stance towards developers who need to store data. Second, and more importantly, Twitter is simply a different online community than Facebook, allowing me to test generalization. (Also, Facebook and Twitter arose from very different communities. Facebook started with college students; Twitter started with early adopters, particularly in and around Silicon Valley.)

Like Facebook, Twitter's service revolves around social networks. Twitter uses variations on the word "follow" to express its network to users. You sign up to follow people; other people sign up to follow you. Unlike Facebook, Twitter network ties are asymmetric: I need not follow you for you to follow me. For instance, anyone can follow President Barack Obama. He will not follow most people back. Twitter limits status updates, called "tweets," to 140 characters. When logged in, you see a stream of tweets by the people you follow (Figure 16). Other clients, like the Tweetdeck client in Figure 17, provide slightly different visual representations of the stream. Users can "retweet" anyone's tweets they find interesting, forwarding the tweet to their own network. Witty comments and web content often spread this way.

What's happening?		140	eegilbert 836 tweets	
			115 234 18 following followers list	ed
Latest: @yapchaian The VC guy? Curious, v @wemeddle? about 3 hours ago	what prompted the	Tweet	Cad • mus n. a tool to find the m relevant tweets in you timeline.	ost r
			Home	
chi2010 Yes. Follow @HC @vinuthap: @chi2010 car wee crr 15 minutes ago via web	LiforPeace for more info RT n any one get a peace badge?		@eegilbert Direct Messages	10
sharoda Where are #chi2 26 minutes ago via Twitterrific	010 people? Drinks anyone?		Favorites Retweets	
sunahsuh If I had a bike	I'd take it out for ride now to	clear	Search	0
drive you nuts. 30 minutes ago via Digsby	d wallow. Living alone really (	loes	Lists outercircle	(
hansent @chi2010 do badges at registration to	on't forget to pick up your pea become an HCI for peace	ace	uiuc 🗎	
about 11 hours ago via Tweetie Retweeted by chi2010			easytomiss 🗎	
yapchaian The local tap tokens for comp drinks. T bottle (cont) http://tl.gd/	rotates objects that function a Tonight's tokens include cork /r2rac	as s and	innercircle A @nancybaym/ internetresearchers	

Figure 16. The popular microblogging service Twitter (http://twitter.com) in 2010. Users of the service "follow" others users in a asymmetric way (i.e., reciprocation not necessary) and their status updates appear all together in a stream such as this one. The stream allows for passive social awareness. From a consumption perspective, the collapsed context problem means that people from various parts of your life are all squished together in one place, the single stream.



Figure 17. The popular Tweetdeck client offers users another way to interact with Twitter. Roughly 10% of Twitter users use Tweetdeck; at least twice as many use twitter.com. It is, however, based on the same principles as the Twitter web interface. Tweetdeck has a columnar layout, with most columns replicating existing Twitter pages, such as "Mentions." The leftmost column, for example, is essentially the same as twitter.com. We Meddle Lists, introduced shortly, can plug into existing tools like Tweetdeck.

#### THE COLLAPSED CONTEXT PROBLEM

This chapter attacks the collapsed context problem. You can think of the collapsed context problem this way: imagine living your whole life at your own wedding. Everyone you know from various parts of your life is there: grandmothers, in-laws, cousins, coworkers, childhood friends, etc. Producing content for a social media site today is like forgetting you have the microphone on: everyone at the wedding hears everything. Consuming content (e.g., reading the stream or the Newsfeed) is very much like standing in the receiving line: everyone you know passes by in random order.

This is, more or less, the state of social media today. As you might imagine, it has problems. First off, everybody sees the same you. Are you the same person at work, at home and in public? danah boyd has termed this the *collapsing of context* (boyd 2002), mostly writing about it from the perspective of self-presentation. Do you want to share your latest party pictures with everybody, including people you didn't invite? But it has other consequence too. In social streams, all these wildly different people come to you through one channel, in temporal order, with nothing distinguishing one from any other.

Want to monopolize someone's Twitter stream? Just write about what you're eating, seeing or doing every ten minutes. Because of collapsed context, your messages will swamp other messages and occupy more attention. In Twitter, this might mean that a movie star's incessant tweeting obscures the one gem a week from your best friend. In the real world, and even with varied media (Haythornthwaite & Wellman 1998), we can easily enforce boundaries: turn on the TV to hear about the movie star; pick up the phone to talk to your best friend. Today's social streams make this much harder.

#### WE MAY AS WELL CALL IT "TEMPORAL MEDIA"

The collapsed context problem has a long history in social media. Although I reference the contemporary service Twitter, the collapsed context problem is present in just about every social technology the internet has seen. I think it is symptomatic of our reliance on time as the central design axis, the thing around which the entire interface revolves. Look at email, instant messaging, IRC, USENET, even the old UNIX tool talk, and you see time as the central design element. You never see relationships, the heart of social media, organizing the design. We may think of it and call it "social media," but it's not much of a stretch to instead call it "temporal media." You never see social media rendered *socially*.

Computationally, we do not understand the relationships people express in social media. I in no way want to minimize or dismiss great qualitative work in this space. However, if we do not understand relationships *computationally*, we cannot architect systems around them. We need to scale to 100 million, 200 million, 500 million users. It seems intuitive to render social media *socially*, but the central challenge is understanding relationships computationally. Existing work has not given us that.

#### TWITTER AND TIE STRENGTH

When a user first signs into We Meddle, the system presented in this chapter, an agent uses the Twitter API and the user's history to compute tie strengths for every person the user follows. The agent applies the model from Chapter 3. Migrating the model to Twitter required careful mapping, but the majority of it remains the same. Table 9 shows precisely how the Facebook model from Chapter 3 became a Twitter model. With the exception of *Participant's number of apps*, the model relocates rather directly. (In fact, this was an explicit goal in the initial work in Chapter 3: identify variables that draw on the breadth of Facebook, but have analogs elsewhere.) One exception is Chapter 3's important Educational difference variable. Twitter does not have this data. (On the whole, Twitter is a much leaner medium than Facebook; participants almost never report these data, and Twitter has no official field for it.) Without *Educational distance*, the Twitter model would not have a single Social Distance variable. To bring social distance into the model, I substitute the log of the difference in follower counts, a "fame differential." Rather than copy the coefficient directly from Educational difference in Chapter 3, Logged follower difference takes as its coefficient the average of all social distance variables: Educational difference, Political difference, Age difference and Occupational difference. In every other case, I purposely and naively dropped the coefficients from Chapter 3 right onto Twitter. Due to this direct, unfiltered mapping, I can ask the following question: How well does the original model generalize to another community, particularly one where intuition might suggest very different dynamics?

Facebook Variables	Twitter Variables	β
Days since last communication	Days since last communication	-0.587

Days since first communication	Days since first communication	0.581
Intimacy × Structural	Intimacy × Structural <sup>†</sup>	0.308
Wall words exchanged	@-reply words exchanged	0.23
Mean strength of mutual friends	Mean strength of mutual friends	0.198
Educational difference	Logged follower difference (fame)	-0.123
Structural × Structural	Structural × Structural <sup>‡</sup>	0.15
Reciprocal Serv. × Reciprocal Serv.	Links exchanged × Links exchanged	-0.146
Participant-initiated wall posts	User-initiated @-replies	0.112
Inbox thread depth	Direct messages	-0.105
Participant's number of friends	User's following count	-0.105
Social Distance × Structural	Social Distance × Structural §	0.1
Participant's number of apps	no analog	
Wall intimacy words	@-reply intimacy words	0.085

Table 9. The top predictive variables as measured by standardized beta coefficients for the *How strong*? model in Chapter 3 and their Twitter analogs. We Meddle uses these features to build models of tie strength in Twitter. *Logged follower difference* substitutes for the original social distance variable, *Education difference*. It captures a difference in fame on Twitter. It's coefficient is decreased from the original because *Logged follower difference* is a substitute: the -0.123 is instead an average of all the social distance variables in Chapter 3. *Days since first communication* also factors in the ordinal value in which the user followed another user.

<sup>†</sup> The *Intimacy* category here is an agglomeration of many variables, an index. It blends two intimacy word measures, reciprocal following relationships, number of mutual following relationships and *Days since first communication*. *Structural* is *Median strength of mutual friends*, as in the original study.

<sup>‡</sup> Both *Structural* variables refer to *Median strength of mutual friends*.

<sup>§</sup> *Structural* again refers to *Median strength of mutual friends*, while *Social Distance* refers to the fame differential variable on line 6.

In part, I designed We Meddle to answer this question. If We Meddle mispredicts a strong tie as a weak tie, or vice versa, users can correct its judgement. (More on this later.) The model in Table 9 is recursive in the same way as the one presented in Chapter 3: it essentially applies ego-centric PageRank. The Twitter model incorporates the 14 variables (or features) listed in Table 9, mixing them in those ratios. But it also incorporates some variables not listed in the rows of the table. (They are listed in the caption). *Days since first communication* not only accounts for the first day of communication but also uses the order in which a user followed someone. *Days since first communication* takes the minimum of these two (standardized) values. For example, if I followed you  $3^{rd}$  in my list of 200 people, but only communicated with you very recently, *Days since first communication* is  $[(200 - 3) - \mu]/\sigma = 1.67$ , where  $\mu$  and  $\sigma$  are the mean and standard deviation of the set {1, 2, ..., 200}, 100.5 and 57.88 respectively.) In total, the Twitter model uses 19 relational features to compute tie strength.

#### WE MEDDLE

*We Meddle* is a web application which uses this model to automatically infer tie strengths between you and all the Twitter users you follow (Figure 18). It calculates tie strength automatically in the background, when users first log in. Users can experience We Meddle in two different ways: *We Meddle Lists* and *We Meddle's client*. (I will explain each in detail shortly, but quickly preview them now.) We Meddle Lists automatically generates up to eights lists, including (most importantly) lists for strong ties and weak ties. In We Meddle's client, users get a richer experience architected around tie strength. They can emphasize strong ties or weak ties, depending on their viewing preferences at the moment. Extending the work of (Fisher et al. 2006), We Meddle uses social history to render its interface. Because different types of people provide different services to us (Constant et al. 1996; Granovetter 1974; Schaefer et al. 1981; Wellman & Wortley 1990), We Meddle's client lets users do things like "only show me weak ties who posted a link" and "emphasize strong ties who said something positive." In a few months on the web, and with only minimal word-of-mouth advertising, over 1,300 people from more than 20 countries have used it.

we meddle'

We build Twitter lists and streams. Yes, we meddle. Like a sweet, nosy grandma. We learn from your past to render Twitter anew.

Sign in with Twitter

You'll come back after you log in.

demo video contact

Figure 18. The homepage of http://wemeddle.com. This is the screen a user sees before she grants We Meddle access to her Twitter credentials. The page gives the user an opportunity to watch a demo video I made for new users by clicking the "demo video" link in the bottom left corner. The video is (currently) available at http://wemeddle.com/wemeddle.mov.

#### Origins of We Meddle

Next, I describe the two core parts of We Meddle, Lists and the client. But first, it seems worth noting the origins of We Meddle. Mozilla Labs<sup>11</sup>, the experimental arm of the organization behind Firefox,

<sup>11</sup> http://labs.mozilla.org

approached us in the Summer of 2009 about building tie strength into Firefox (via an add-on). In short, the idea was the following. Firefox is with you when you visit GMail, Facebook, Twitter, etc. What if it could make sense of the data it sees there? Tie strength seemed like an essential component to understanding the relationships formed in online social media. Ideally, any developer could then build a modification to a site like Twitter (using a Firefox-modding system) that leverages the social store of data. For example: only notify me of new GMail when a message arrives from a strong tie.

After meeting with people at Mozilla Labs HQ, I started building the infrastructure. I built a Facebook data observer (I called them *observers*) and an alpha Twitter observer. However, midway through the project, Mozilla Labs had to move its development staff to other priorities. So I suspended work on the project. However, it seems like a truly helpful tool, and one with lots of thorny issues, like privacy and security of the data store. I encourage anyone who follows to explore this area, as it might tell us something important about relationships across media—something we understand very little about.

#### WE MEDDLE LISTS

We Meddle has two main parts: We Meddle Lists and We Meddle's client. After logging into We Meddle for the first time (Figure 19), a user sees We Meddle Lists. By clicking a link in the top-right corner of the page, she can go the client. (More on the client in the next section.) We Meddle Lists creates up to eight lists automatically for the user: *Inner Circle* (i.e., strong ties), *Outer Circle* (i.e., weak ties), up to four automatically discovered communities (in the social network sense), *Infrequent Tweeters* and *Eager Tweeters*. In Figure 20, you see my strong ties, weak ties and two communities roughly corresponding to SIGCHI researchers and Harvard's Berkman Center. Figure 21 shows a close-up view of my Inner Circle list with two users dropped.

## we meddle

We build Twitter lists and streams. Yes, we meddle. Like a sweet, nosy grandma. We learn from your past to render Twitter anew.



Figure 19. The We Meddle wait screen during a user's first login. Clouds float by on screen as We Meddle digs through the user's Twitter history. The text below says "We're chatting with the ghost of Twitter past. It might take a minute. He's old, you know." After the computation completes (mostly I/O bound), the user gets redirected to We Meddle Lists, illustrated below.

In late 2009, Twitter released a new feature built directly into its infrastructure, called Lists<sup>12</sup>. Lists let users group accounts together under one name. It seems clear, from the company's public statements, that Twitter intended Lists as a way for users to curate Twitter. For example, I could make a list for "economics," curating that part of Twitter for other people. From discussions on the web, and chats with other users I knew, it seemed clear to me that no one used Lists to make their own, personallymeaningful groups. It just wasn't worth the effort. Why put all the work into creating a list no one else would see? That does not mean that people would find no benefit in grouping accounts they follow, but that they did not want to put in the effort to make them.

Lists seemed like a perfect opportunity to test computing tie strength in Twitter. At the time of the Lists release, I was working on the client I'll introduce later. It was an exciting idea, but clearly one that would require buy-in from users: they would have to abandon their existing practices, their existing clients, and use We Meddle entirely. Lists seemed like a perfect opportunity for the tie strength model: automatically make tie strength lists for users with the option to tell system what it got wrong. It would be much easier to subtract a few mislabeled accounts than to add each one individually. We Meddle Lists would give people a way to use tie strength without requiring buy-in.

<sup>&</sup>lt;sup>12</sup> http://blog.twitter.com/2009/10/theres-list-for-that.html



Figure 20. The We Meddle Lists web interface, grouping a user's Twitter contacts with an underlying computational model of tie strength. Upon sign-in, We Meddle automatically computes tie strength for everyone the user follows. The two groups on the left (e.g., Inner Circle and Outer Circle) result from the tie strength computation. The two groups on the right come from community detection on the underlying social network. (See text for details on its construction.)



Figure 21. The author's inner circle, but this time I have dropped two accounts from the automatic suggestions before I make the list. We Meddle suggests people who belong as strong ties, but invites users to make corrections, under the assumption that it's easier to subtract a few accounts rather than add each one individually. Plus, no one seemed to be serving Lists, an opportunity to "fix a problem; get some data." So I hurried to produce an application that used tie strength to make Twitter Lists, putting the client on hold to make it. In January 2009, I rolled We Meddle Lists out to the web. Anyone with a Twitter account could (and currently *can*) try it. The main page (Figure 18) also has a demo video<sup>13</sup> so potential users can see what it does before they log in.

The engine behind We Meddle relies on the Twitter API to both compute tie strength and store lists. After We Meddle renders its lists for the users, she can drop an account by clicking on it. (A note at the bottom left of the interface reminds her that she can.) When she clicks, the profile picture goes to 25% opacity, holding its place to remind the user that she dropped the account. When she's happy with the contents of the list, she can create it, storing the list in Twitter. (She can also change a list's name by clicking on the label, creating a text area.) Storing the list in Twitter means that she can access it from any Twitter client, using it in ways that particular client affords. For instance, the popular Seesmic web client<sup>14</sup> lets users view each list in its own column, meaning that you can slice the conventional Twitter stream into multiple views. A We Meddle Lists user could go back to Seesmic and see their Inner Circle flowing in one column, their Outer Circle flowing in another and a particular social community (e.g., "CHI research") flowing in yet another. (The demo video on the main We Meddle site shows how this could look.) It is a direct attack on the collapsed context problem.



vitak shared this idea 3 months ago.

## List adds

I love this idea. You got most connections dead on, but a couple were in the wrong list. Why not add a feature so you can not just remove a person from a list but also add someone to a list?

Figure 22. A We Meddle Lists user who left feedback about the site on We Meddle's accompanying online community. She's happy with the concept, but wonders why We Meddle does not allow list additions. I explain this decision in the text, but in short: I constrained the interface to better ensure I could study tie strength.

Users cannot add people to the lists We Meddle auto-generates. They can only drop them. This is by design. A few people did explicitly ask for this feature in We Meddle's online forums (Figure 22 shows one such request), however I specifically chose to leave this feature out of We Meddle. Adding accounts to the We Meddle lists could change their meaning. How can we be sure that *Inner Circle* still means strong ties? Perhaps the user started from We Meddle's suggestions but branched off to create a list with a slightly different meaning. This is not inherently bad. If I wanted to study which lists people want to make, rather than computational tie strength, I would allow users to add anyone. But, I want to study

<sup>13</sup> http://wemeddle.com/wemeddle.mov

<sup>14</sup> http://seesmic.com/web

tie strength. Limiting users to deletions (at least inside the interface) is a crucial experimental decision. It allows me to argue that Inner Circle and Outer Circle still retain their tie strength meanings. When a user removes an account from the Inner or Outer Circle, we learn where the model makes mistakes. These clicks are crucial experimental data, as I can draw conclusions about the computational tie strength's generality and its Twitter subtleties. From the click data, I can retrain the model, learning which features to weight differently. When a user drops an account from a list, they effectively relabel a strong tie as a weak tie, or vice versa. In other words, during the natural process of using We Meddle, users leave a trail from which I can study tie strength. Of course, users can add people to their Inner and Outer Circle lists outside We Meddle. They can go back to twitter.com or anywhere else that has a list interface and put whoever they want in the list. I take a closer look at the practice of adding people afterward in the follow-up interviews presented in this chapter.

From a social science point of view, we care most about the Inner Circle and Outer Circle lists. They give us tie strength data. I created the social communities lists to give users more value in the site. In other words, I thought users might not come to We Meddle for two lists, but maybe they would come for eight. More importantly, I offer users the ability to convolve tie strength with community in the We Meddle client. Offering it in the lists interface made for no more work, but gave users more value. I built the social communities part of We Meddle from a freely available network clustering package, called MCL (Dongen 2000). It does a random walk of a network, noting that nodes within a community will have more paths between one another than nodes in different communities. It is by no means the most efficient algorithm for optimizing modularity (Newman 2006), but the package offered an open source implementation in C— parameterizable, tweak-able code that runs quite fast on networks of We Meddle's size. We Meddle's network I/O dwarfs MCL's running time, so it adds little to a user's overall wait time. I tuned the parameters for community detection based on pilot studies on Twitter networks in our lab. We Meddle Lists displays up to four of these automatically detected communities. A list must have five accounts (an arbitrary, but useful cutoff) in it for We Meddle to display the list to a user. Lists also generates *Eager Tweeters* and *Infrequent Tweeters*, primarily with the intention of using the lists in We Meddle's client.

#### WE MEDDLE'S CLIENT

If you buy into the We Meddle way of life, you can also use its full-featured Twitter client. The We Meddle Twitter client, like other Twitter clients such as Seesmic or Tweetdeck, shows users the stream of tweets by people they follow. However, the We Meddle client makes one very crucial break from the dominant design paradigm of social streams: it does not treat everyone equally. Different people get different amounts of screen real estate. Twitter and Facebook, on the other hand, give equal prominence to everyone. The client's main slider control (Figure 23, left), a reference to "inner circle vs. outer

circle," gives users the ability to emphasize strong ties at the expense of weaker ones, or the other way around. As a user slides toward the inner circle or the outer circle, We Meddle adjusts profile picture size, typeface, font size and opacity. Pushing the slider all the way to the Inner Circle totally removes weak ties from the interface, and vice versa. Placing the slider in the middle essentially replicates twitter.com, where everybody appears the same. The slider remains in place across page loads and log-ins to support stable mental and visual models of the interface. I like to think of it as a *socially zooming interface* (Bederson & Hollan 1994). (The interface updates in real time, using Javascript.) Figure 23 shows a screenshot of me using the We Meddle client to render my own Twitter stream. In it, I have used the control at left to skew the interface toward my strong ties, at the expense of my weaker ones. Strong ties appear larger, in a more demanding font and with more saturation (i.e., they demand more visual attention).



Figure 23. The We Meddle Client web interface, rendering a user's Twitter stream using an underlying computational model of tie strength. Here, a user has moved the slider slightly toward the inner circle to emphasize strong ties at the expense of weaker ones. We Meddle translates emphasis into profile image size, typeface, font size and opacity. Users can also apply filters (upper right) to reduce the stream. For example, with the interface still rendered by tie strength, users can filter the stream to only particular social communities, like web researchers.

The We Meddle client takes its inspiration from real life social relationships. In real life, we do not pay everyone in our social lives equal attention. Imagine the following scenario: somebody you just

bumped into once at a conference starts calling you constantly. You don't have caller ID; you either pick up the phone blindly or just let it ring, potentially missing calls you care about. You have to listen to what they say before you can free up the phone for the next call. The behavior is totally out of line. We would be furious with the guy. But, we permit a very similar thing from our stream clients. Everyone gets unlimited access to your visual attention, as much as they like. We Meddle tries to make a guess about your social history with the people you follow, and then renders the interface accordingly completely on the client end.

00	We Meddle Client.	$\bigcirc$
C	http://wemeddle.com/client/	☆ Q- Google
we	meddle	TWEET! FEEDBACK   LOGOUT
6	Error rates have improved significantly, but remain well above acceptable levels. Working. AJK — TWITTERAPI 13M AGO PRX is biring! Great dia. Join award winning tech team @PRX creating web & mobile apos	communities
<b>7</b> 41	for pub media http://ow.ly/1WgOA (via @jakeshapiro) — BERKMANCENTER 53M AGO	show only UIUC KIDS 9 show only WEB RESEARCH 3
	I adore the @foodcoop and have been a happy member for nearly 5 years, but I hate how life always gets in the way. Make-ups, here I come! — WAITINGFORPOPPY 1HR AGO	show only CHI RESEARCH 36
*	A pretty good World Cup Google Calendar http://bit.ly/8hU2mb	more meddling show only POSITIVE TWEETS 12 show only NEGATIVE TWEETS 3
91	@zephoria It's not unusual for successful business women to be conservative in all kinds of ways. It's part of what made them successful :-( $-$ NANCYBAYM 3HR AGO	show only FREQUENT POSTERS 21 show only INFREQUENT POSTERS 14
ilt	Just bought the Pulse newsreader app for iPad. Impressed! http://www.alphonsolabs.com/ by @pulsepad — ILOVETYPOGRAPHY 4HR AGO	snow original view
	RT @FeministBreeder: I was up late creating a new web project dedicated to doulas. Check it out and submit your photo to the project! ht $-$ waitingforpoppy shr ago	
<b>91</b>	It's so easy to get inspired to go to work when you know boyd, Nakamura, Sandvig, Monroy Hernandez & other awesomeness awaits you there. — NANCYBAYM SHR AGO	
	RT @radiofreegeorgy: @vrhinesmith I'm happy to be there to balance the scales. [@rww: look forward to seeing you there Georgy!] $ _{\rm RWW  GHR  AGO}$	
	Web Apps Are One Thing But What About Security on the Smart Grid? http://bit.ly/cnQl3g — RWW 12HR AGO	
	RT @dlifson: @nww I bought my ticket! See you there. You guys were the first to write about us (@postling) 9 months ago, so we love you. — RWW 13HR AGO	
	I enjoyed the #cmsummit these past 2 days, quality speakers and I learned a few things! Congrats @johnbattelle, @janerri and team. #IWNY — RWW 14HR AGO	

# Figure 24. The We Meddle Client filtering tweets to only show ones it thinks are positive. (This is a different stream than the one above.) We Meddle uses a set of word-matching heuristics borrowed from previous literature to mark tweets as positive or negative. See text for details.

This is the client's main feature: a stable rendering of your Twitter stream according to automaticallyinferred tie strength. (The client interface also has a way for users to override the system's judgement. By hovering over a profile picture, a user can tell We Meddle that it got someone wrong.) As you socially scale the interface, however, it does not naively scale all the parameters individually and at the same rate. I worked hard to make the interface legible while at the same time placing emphasis on whatever the user chooses. Figure 23 shows the interface one tick away from equal; "equal" meaning essentially a replica of the standard twitter.com stream interface. When you move two ticks away from center, (in the case of moving towards inner circle) the strong ties get bigger and the weak ties get smaller. At three ticks from center, the weak tie profile pictures become grayscale, demanding even less visual attention. (When I've been away from Twitter for a few days, I often use this setting to quickly scan backward.) At four ticks from center, the weak ties disappear entirely from the interface, leaving only strong ties. (This example focuses on strong ties, but users can apply it as easily to weak ties.) Users can also easily do normal things that all clients provide: they can tweet, reply to a tweet and retweet. When building We Meddle, I found myself constantly walking the fine line between building out the research part and the building enough expected functionality so that people would stay and use it.

#### Communities

People talk about their social networks in two primary ways: tie strength and communities. While this dissertation is about tie strength, We Meddle tries to support this practice in full, integrating it into the way people can consume their stream. As described above, We Meddle uses a community detection algorithm to decompose a user's following graph into up to four relatively distinct social communities. Figures 23 and 24 (upper right) show my four communities, "Design," "UIUC Kids," "CHI Research" and "Web Research." I gave the communities those names using a auxiliary-interface, allowing users to see who belongs to a community before they give it a name.

At any time, meaning under any tie strength skew, users can click on a community name to filter the stream to only that community. This means that users can compose tie strength and community structure. For instance, I can tell We Meddle to show me only CHI researchers, but emphasized by tie strength. Or, while I'm viewing CHI researchers, skew them the other way towards weak ties. The transitions between communities and the full stream happen via a smooth animation using the jQuery<sup>15</sup> Javascript library. The filtered-out tweets quickly fade to white, but hold their space, and then slide up—giving a smooth transition between views. This is the first interface I am aware of to allow users to view their streams they way they talk about them.

#### More meddling

From social support (Schaefer et al. 1981; Wellman & Wortley 1990) to finding new information (Constant et al. 1996; Granovetter 1974), the literature strongly and repeatedly suggests that we get different things from different kinds of ties. To support this in Twitter's mediated environment, We Meddle provides a set of filters: Links, Positive Tweets, Negative Tweets, Frequent Posters and Infrequent Posters (see upper-right of Figure 24). The positive and negative sentiment filters borrow data and techniques from (Pang et al. 2002) and LIWC (Pennebaker & Francis 1999). As with social zooming, the filters persist until the user unsets them. (When a user applies a filter, the same smooth animation takes them to the stream's new state.) Used in conjunction with social zooming, users can powerfully filter tweets,

<sup>&</sup>lt;sup>15</sup> http://jquery.com

i.e., "only show me tweets by weak ties who post infrequently" or "emphasize tweets by strong ties with negative sentiment from the CHI community." In the latter case, imagine finding a friend who needs a pick-me-up.

The two sentiment filters do not always produce the right answer because they only word-match. For example, the sarcastic comment "Traffic is just freaking lovely today. I'm going to miss my meeting. Great!" will utterly confuse We Meddle's sentiment filters. To compensate for this problem, and since the focus of this dissertation is not sentiment, I tuned the filters rather conservatively. They each need to see at least three decidedly positive or negative words in a tweet to declare the whole tweet positive or negative. The infrequent and frequent filters filter the stream based on a person's usage of the service. In building them, I envisioned use cases where someone wants to make sure they do not miss a rare tweet by someone close to them. (Followup work might consider allowing filter inversions too: filter out anyone who tweets *a lot*.)

#### Breaking the temporal model

After the brief Mozilla experiment described earlier, I built a prototype We Meddle client that broke the temporal model of social streams. By "broke," I mean that it no longer showed the tweets ordered by time; it split the screen into two panels, one for strong ties and one for weak ties. Strong ties occupied more visual space, and the user could expand or contract each group using sliders.

Early response to this prototype from the users on the web was more mixed than the final client available now. People simply missed time-ordering. Many people expressed (either through tweets or private email to me) that they had grown so accustomed to streams strictly ordered by time that We Meddle required too much cognitive effort to read. I took this as a failed experiment and moved onto a more gentle bending of the temporal model, the client's current incarnation.

#### DEPLOYMENT

Beginning in early January 2010, I made We Meddle public and open to Twitter users on the web. At first, I limited usage to by-invitation-only: existing users could invite up to five new users, with a hand-ful initially seeded by email. In late January 2010, I lifted this restriction and allowed any Twitter user on the web to use We Meddle. (I lifted the requirement after I had worked out early bugs and could guarantee a reasonable level of service and responsiveness.) We Meddle was (and currently is) available on the web at http://wemeddle.com.

I could have done a lab study: recruit Twitter users from around campus to log in and tell us whether the tie strength calculations match up with their thoughts on the people in their network. But two crucial points make this field study a better, more reliable method. First, a real life field study helps establish if this approach helps anybody solve any problems. Do people come to use it? What do they think about it? Second, by lifting the college campus sampling frame, we can be more confident in the analytic results that come from a field study. We Meddle has seen users from all over the globe—see the Google Analytics<sup>16</sup> heatmap in Figure 25. (This map has interesting implications for the linguistic analysis components of We Meddle; the sentiment engine fails for non-English tweets.)



Figure 25. A heatmap of We Meddle's user distribution across the world. Most users come from the United States, but China also has a strong presence, accounting for between 6% and 12% of users (depending on whether you measure in visits or raw page views). The maximum on the legend in the lower-left corner is 2,403. (Image generated by Google Analytics.)

To date, over 1,300 people have logged into We Meddle, with minimal press and advertising. Many more have visited the site, but since it requires a login and read and write access to your Twitter account, I suspect many turned away at the login prompt. One difference between a controlled lab study and a field study on the open web is that anyone can (and will) respond to your application. They do it not because you asked them to do it, but because they want to talk about it. At the end of this chapter, I show many of the unprompted reactions from Twitter and the web. (Because Twitter only allows searching tweets from the last nine days, these tweets are now hard to locate). By and large, the response from users was overwhelming positive. I also built an online community forum into the site using the service Get Satisfaction<sup>17</sup>, a popular and emerging web standard for such things. Users can leave feedback there as well, but most feedback came from Twitter itself.

<sup>&</sup>lt;sup>16</sup> http://google.com/analytics

<sup>17</sup> http://getsatisfaction.com

Of course, we would expect that people who dislike We Meddle might not speak up as often as people who like it. A follow-up interview study presented later in this chapter uses random sampling to try to uncover the views of people who did not speak up. More broadly, We Meddle represents the kind of research I hope the community uses more often. Make a service that solves a problem that anyone can actually use. It has the upside that we can do away with representative sampling problems: if someone comes of their own free well, uses the service and it helps them, we're done. IBM Research calls this approach *Venture Research* (Greif 2009; Lau 2009). (They have also carried out one of its largest examples to date (Viégas et al. 2007).) What better way to show that users find value in a system than when they use it of their own free will?

#### Architecture

I would be remiss in a Computer Science dissertation not to mention the actual system implementation details. Figure 26 shows a high-level overview. I wrote the core We Meddle tie strength engine in Perl, appropriating the output of an R<sup>18</sup> statistical model. I chose Perl for its speed, its ease with text and the Twitter library available for it on CPAN<sup>19</sup>. When a user first signs into We Meddle, it needs to build a database of tie strengths associated with each account the user follows. The sign in forks off hundreds and sometimes thousands of API requests against Twitter. This was the main technical hurdle: overlapping the API requests in precisely the right way with many simultaneous users logged in. It took tweaking and hacking, but finally worked.

<sup>18</sup> http://www.r-project.org

<sup>19</sup> http://cpan.org



Figure 26. The architecture of We Meddle. When a user signs in for the first time, a process forks off hundreds and sometimes thousands of requests against the Twitter API. As they come back, they filter through the simple application of the tie strength model and then percolate through the user's ego network. Usually in under one minute, We Meddle generates its lists from the output of the model.

As the API requests come back, they filter through the non-network (non-structural) parts of the tie strength model. After they all come back (many overlap), the Perl-based model percolates the non-structural tie strengths through the user's ego-network. I could solve this with an eigenvector centrality algorithm, but in practice a simple percolation loop with five iterations seems to always converge (and is *much* faster and consumes fewer resources). Near the end of the tie strength computation, I project all tie strengths onto a [0, 1] interval. Anyone one-half standard deviation above the mean gets marked as a *strong tie*; anyone two-thirds of one standard deviation below the mean gets marked as a *weak tie*. One and one-half standard deviations earns you a *very strong tie* designation; two standard deviations below the mean earn you a *very weak tie* designation. I found these cutoffs through trial and error during pilot tests in our lab, in an attempt to reduce noise and engender trust in the system.

After computing all tie strengths, they get written to a MySQL<sup>20</sup> database where We Meddle can read and act on them. The web interfaces seen by We Meddle users are written in PHP. The client side code is written in Javascript, making heavy use of the jQuery<sup>21</sup> toolkit for animations and asynchronous

<sup>&</sup>lt;sup>20</sup> http://mysql.com

<sup>&</sup>lt;sup>21</sup> http://jquery.com

communication. The server side of We Meddle interacts with Twitter via an open source Twitter library called *twitter-async*<sup>22</sup> that I modified to serve We Meddle's specific needs.

#### DOES IT GENERALIZE?

When a users tells We Meddle that it got something wrong in either interface, we can consider it a mistake by the model. How often does this happen? I think it's fair to say that the model generalizes when we see a comparable number of mistakes to the number we saw in Chapter 3. In other words, it generalizes to Twitter when roughly 11-12% of its decisions are mistakes. Even so, I should caution that this would only say that the model generalizes to *one other* online community. We cannot say that it generalizes to email or instant messaging, for example. However, we have almost no data on generalization in online communities research, so this would be a big step.

We Meddle received 1,105 corrections from users. Again, this means that a user either removed someone from a We Meddle-generated list or they corrected a someone's tie strength in the client. These 1,105 corrections came from 236 different users. So, most We Meddle users made no corrections. We could view this as a huge success: the overwhelming majority of users experienced complete and utter success with We Meddle! Of course, this seems blindly optimistic. Some people who made zero corrections probably thought We Meddle got everybody right. (I have some qualitative data to back this up.) However, probably more did not realize they could correct We Meddle's predictions (qualitative data also supports this), or felt it wasn't worth the extra effort to make the clicks. I have to deal with these realities because We Meddle is not a controlled experiment. I did not tell users to correct every mistake. I let them do what they wanted. (The interfaces try to make it clear that they *can* correct We Meddle's judgements. The demo video also makes this clear, accessible from many places on the site.)

Here, I will pursue a conservative estimate, likely an upper bound on We Meddle's error rate. First, we will limit ourselves to only the 236 We Meddle users who made at least one correction and who made at least one list, leaving out the others who did not. Note that this procedure leaves out anyone who felt We Meddle's predictions so closely matched reality they did not correct the system. But we know that this group of users understood the process by which they could correct We Meddle. All 1,105 corrections came from these 236 users. From these users, We Meddle saw 27,529 relationships for which it estimated a tie strength. Of these, users only had a chance to correct accounts that We Meddle marked as a strong ties or weak ties (see above for the specific method), a total of 14,075. So, users corrected 1,105 out of 14,075 potential accounts, or 7.85%. However, as noted above, users had no way to tell the system that it had left someone out: they could only drop users from lists. Since We Meddle markes approximately equal mistakes in each direction (see analysis in Chapter 3), we can simply double this

<sup>&</sup>lt;sup>22</sup> http://wiki.github.com/jmathai/twitter-async

percentage, obtaining an upper bound of 15.7%. While slightly higher, it closely resembles the expected error rate of 11-12%, especially since 15.7% is a conservative upper bound. (Note that a Chi-square test here is inappropriate due to the approximate nature of We Meddle's data versus the certain data in Chapter 3.) From this number, I conclude that computational tie strength generalizes to at least one other online community with almost no modifications. In other words, I make the cautious claim that some general features of online communities and relationships in those communities transcend the bounds of particular implementations on particular sites.

#### Mistakes in terms of predictors

Next, I study how the corrections to We Meddle's predictions express themselves in terms of the model's input predictors. Tables 10 and 11 summarize these analyses. They tell us how a tie strength model for Twitter differs in subtle ways from the tie strength model built from Facebook, and how we might alter it to learn a slightly better one. They also tell us which fundamental changes to the model could end up as most profitable. The two tables show We Meddle errors broken down by type: strong tie mistakes and weak tie mistakes. Again, I compare only data from those 236 users who made corrections. Each table compares two groups: correct (strong or weak) predictions and incorrect (strong or weak) predictions. As the lists were split in the interface, it seems reasonable to split them in the followup analysis, as users most likely experienced them this way. To compare the groups, I perform within-user standardization. That is, I compare scores standardized against all relationships for a particular user. As the same model generated both the correct and the incorrect predictions, any difference stems from the difference between the populations, not some artifact of the model. We Meddle users made many more corrections to strong ties than they did to the weak ties, with 79.3% of all user corrections coming from strong ties. (This is to be expected, as users probably care more about getting their strong ties exactly right.)

Tables 10 and 11 summarize the analysis of mistakes. (The captions on each table document specific terms in the tables; I will focus on high-level trends here.) The thing that jumps out most is the contrast between strong and weak tie mistakes. Strong tie mistakes exhibit a strong recency effect, with correct predictions 0.07 standard deviations more recent on average than mistakes. (The tables uses medians as measures of central tendency to account for the non-normality in the distributions, even after stan-dardization.) We see no similar effect for weak ties. The most striking difference, however, between strong and weak tie mistakes is the impact of the network in strong tie mistakes. Four network predictors reveal big differences: *Intimacy × Structural* (0.263 standard deviations different), *Mean strength of mutual friends* (0.073 standard deviations different), *Structural × Structural* (0.134 standard deviations different) and *Belongs to community* (a non-model, categorical feature). True strong ties also have lower user id numbers (indicating longer Twitter membership). But the network features pack the biggest

punch. Interestingly, *Belongs to community* may be the most instructive. It refers to membership in one of the four biggest communities detected for interface reasons in both Lists and the client, and it tells an interesting story about perhaps the biggest vulnerability of the tie strength model. (I never intended to analyze ties *in terms* of it, but I had the data as a consequence of design decisions.) Membership in one of the four big communities (dense network clusters with more connections inside than outside) impacts the model's the network component. Recall that the tie strength model blends tie strength recursively through the ego network: tie strength is a function of the tie strengths of mutual friends. For large clusters, this implies that many people benefit from a strong tie in a large cluster. (Statistics like the mean are especially sensitive to these situations.) As a group, these network features make a strong case for a more refined vision of the network in the next tie strength model.

However, what should the network model look like? In the current version, everybody in the mutual friends ego network contributes equally the network component of the tie strength model. But why should my wife's best friend contribute to the tie strength of my best friend, if they only articulated a friendship because they felt obligated? So I propose a weighted network model, one that reappropriates the one presented in this dissertation, but weights each mutual tie by tie strength. Practically speaking, however, this is easier said than done. To do it, not only would we have to estimate every ego-centric tie (as We Meddle does), but also all of the cross-cutting ties between alters (i.e., the other people in the mutual network). To operate in web response time (as We Meddle does), this is intractable; plus, we simply cannot see everything every pair does. However, what we might consider is a low-fi tie strength using only a handful of tie strength predictors, drawn from roughly orthogonal dimensions.

For example, a better network model might use *Days since last communication* and *Intimacy words* to get a very rough estimate of cross-cutting links in the network of mutual friends. A new model could then use those low-fi tie strength estimates to weight their contribution to the overall tie strength of interest, yielding a more sophisticated and refined model. In fact, we might think of this in terms of weighted community detection: brittle links (i.e., low tie strengths) have a tendency to break during community detection. A model like the one I propose here would have similar dynamics: tie strengths would inherit their values from heavily weighted peers, and lower-valued ones would contribute less. This would of course require deeper knowledge of the communication record (e.g., being inside Twitter) or it would require an extra crawl of the network at computation time. Either way, it costs time. But as we see in Table 10, it's the most profitable place to put extra effort when it comes to bringing down the error rate.

For weak tie mistakes, we see a completely different picture. Weak tie mistakes express themselves most distinctly in *@-reply words exchanged*, a 0.11 standard deviation difference, with correct predictions lower than mistakes. It seems natural to infer from this finding that it's easier to message someone you

do not really know on Twitter than on Facebook. Whereas on Facebook you need to confirm my friend request before we can exchange messages, I can @-reply Oprah Winfrey on Twitter. There is no barrier. Curiously enough, we again see a substantial difference for *Belongs to community*, although less of difference than we saw for strong ties. For strong ties, it seems that the model does not capture network features well enough. For weak ties, we do not see effects for network-based predictors, like *Intimacy* × *Structural*. I do not have a great explanation. This seems like an excellent place for future study.

Tie Strength Predictor	+ median	– median	W or X <sup>2</sup>	р
Days since last communication	-1.252	-1.323	1,376,838	<0.001
Days since first communication	1.125	1.115	1,230,026	0.808
Intimacy × Structural	-0.249	-0.512	1,373,856	<0.001
@-reply words exchanged	0.111	0.082	1,270,100	0.278
Mean strength of mutual friends	0.671	0.598	1,348,726	<0.001
Logged follower difference	-0.168	-0.178	1,288,437	0.091
Structural × Structural	0.715	0.581	1,360,934	<0.001
Links exchanged × Links exchanged	-0.167	-0.140	1,264,614	0.367
User-initiated @-replies	0.329	0.339	1,273,548	0.230
Direct messages	-0.120	-0.146	1,255,016	0.558
User's following count	too little data	a		
Social Distance × Structural	too little data	a		
@-reply intimacy words	0.133	0.039	1,312,502	0.013
Non-model variables				
States personal url			X <sup>2</sup> =0.549	0.473
Is high frequency			$X^2 = 0.671$	0.413
Is low frequency			$X^2 = 1.132$	0.287
Belongs to community			X <sup>2</sup> =37.43	<0.001
User id	-0.257	-0.195	1,138,977	<0.001
Tie strength	1.273	1.285	1,230,481	0.820

Table 10. An analysis of We Meddle's *strong tie* mistakes in terms of the model's input predictors. Also shown at the end of the table are non-model variables collected for various reasons: diagnostics, interface reasons or as the output of the model itself (tie strength). The table compares two groups, correct predictions and incorrect predictons, but limited only to those users who made a correction. + *median* refers to the median among the correct strong tie predictions; – *median* refers to the median among incorrect strong tie predictions. W refers to the Wilcoxon statistic. Note the strong recency effect as indicated by *Days since last communication* and the many effects for network-based predictors: *Intimacy x Structural, Mean strength of mutual friends, Structural x Structural* and *Belongs to community. Belongs to community* here means that the partner (not the user) belonged to one of the biggest four communities generated for interface reasons. A Bonferroni correction for all the tests done here means that  $\alpha = 0.05/18 = 0.00278$ . (All "< 0.001" rows are therefore significant under standard hypothesis testing norms.)

Tie Strength Predictor	+ median	– median	W or X <sup>2</sup>	р
Days since last communication	0.631	0.587	602,953	0.451
Days since first communication	-0.504	-0.420	519,727	0.015
Intimacy × Structural	0.123	0.123	604,179	0.423
@-reply words exchanged	-0.306	-0.195	456,055	<0.001
Mean strength of mutual friends	-0.113	-0.254	620,182	0.157
Logged follower difference	-0.133	-0.163	685,398	<0.001
Structural × Structural	-0.422	-0.503	623,827	0.12
Links exchanged $ imes$ Links exchanged	-0.198	-0.207	609,755	0.310
User-initiated @-replies	-0.342	-0.296	532,777	0.053
Direct messages	-0.215	-0.221	640,661	0.027
User's following count	too little data	1		
Social Distance × Structural	too little data	1		
@-reply intimacy words	-0.303	-0.277	549,013	0.189
Non-model variables				
States personal url			$X^2 = 0.049$	0.824
ls high frequency			$X^2 = 0.096$	0.757
Is low frequency			$X^2 = 1.393$	0.238
Belongs to community			$X^2 = 13.18$	<0.001
User id	-0.170	-0.211	583,004	0.992
Tie strength	-0.793	-0.889	693,174	<0.001

Table 11. An analysis of We Meddle's *weak tie* mistakes in terms of the model's input predictors and non-model variables. The table compares two groups, correct predictions and incorrect predictions, but limited only to those users who made a correction. Table headings refers to same things as the table above. We see very different reasons for weak tie mistakes than we do for strong tie mistakes. The network-based predictors all but vanish, with *Belongs to community* contributing, but in exactly the opposite direction that it does for strong ties. We also see strong effects by *Logged follower difference* and *@-reply words* exchanged, perhaps signaling the ease with which you can message a higher-status user on Twitter. As opposed to the strong tie mistakes, we see that model makes mistakes as a function of tie strength, with mistakes more likely the lower the tie strength.

#### HOW USERS EXPERIENCED WE MEDDLE

The previous section tells us that the tie strength model generalizes, but it does not tell us about users' feelings toward We Meddle, or whether it solved any of their problems. To answer these questions, I couple the quantitative techniques above with qualitative work. In addition to the interviews I report next, I also present many unprompted reactions to We Meddle from Twitter and the web.

To recap, our research question is the following: "Can a computational tie strength model help alleviate design challenges introduced by the collapsed context problem?" Or more simply, does computing tie strength help users or make social media better in some way? I sampled We Meddle users at random from the logs, contacting them by @-mentioning them through Twitter. I then setup semi-structured interviews with eight We Meddle users in whatever medium they preferred (e.g., phone, IM or email). The participants ranged widely in backgrounds and in how they used Twitter, from young coders in the tech industry to small business owners who primarily use Twitter as a promotional tool. The point of the interviews was to elicit feedback on the system and the model from people who did not say any-thing publicly. The interviews, but I engaged in follow-up conversations to clarify certain points. Typically, participants logged into We Meddle during the session to look at the lists while we talked. Finally, the interviews were coded using a grounded-theory influenced approach where themes arise iteratively (Glaser & Strauss 1967).

When I asked about the actual composition of the tie strength lists, most users had very positive things to say, reflecting the once-in-awhile error rate described in the quantitative section.

[Does the Inner Circle reflect your real-life Inner Circle?] Um, pretty close. Yeah, I mean about as close as I can get. There's a few people on there that I communicate to a lot, and I would not define them as being in my Inner Circle, however, mostly yeah.

[Did the lists reflect your real social life?] Um, I was pretty amazed to tell you the truth. Really, amazed cause uh ... The one I had an extreme hard time with trying to figure out was the Outer list. And this was probably the same with most people cause they're not people you communicate with much. So, I only remember one person I actually recognized on the Outer list. But the other three [Inner Circle, communities] were pretty close to right on.

Another interesting thing is that my Outer Circle is composed entirely of people I want to meet. Yeah, it's all the people I'd want to meet. It's really funny. And most of them are tech people like me.

Most users expressed overall satisfaction with the quality of the lists We Meddle produced. (Interestingly, in the last quote about Outer Circle people he wants to meet, the participant's sentiments clearly resonate with the literature on weak ties.) In a few instances, users even relayed their surprise when We Meddle correctly identified certain people (or types of people) in the lives.

A few of the ... well I remember a few of the people in the Inner Circle are actually relatives, and that was pretty cool. I didn't expect that.

It's actually kind of fun to look at the Inner Circle and say, "Wow, look at that person. I haven't talked with him in a *long* time, but they totally fit there." Yeah, there's some of those in here.

It's interesting that it actually placed my girlfriend *four rows down* versus at the very top, where I would expect her to be. [But she's in the list?] She's in the list, yeah absolutely, most of the lists actually. I hope she doesn't see the four rows down part. [participant laughs]

Many users told a variant of the following story: The Inner Circle picks off some of the most important people in various aspects of my life and puts them all together. (We see a very similar told by Ethan Zuckerman in the video presented at the end of this Chapter, in fact.) You see that story reflected in the quotes above. The participant with the girlfriend *four rows down* illustrates a side effect I did not anticipate. Some participants inferred an order on the lists I did not intend. We see something similar in a blog post included at the end of the chapter, where a We Meddle user posted a screenshot of his Inner Circle, apologizing to the few people We Meddle left out.

While participants mostly reacted positively to the construction of the lists, there were problems. I dug at these issues, to see what we can learn about the theory and application of computational tie strength. In some cases, I think more documentation on the site could have resolved some confusion about particularly the Outer Circle. Because users did not often feel very connected to these people, they wondered why We Meddle had even made a list out of them.

I couldn't figure out the Outer Circle, if those are people I just communicate to like once a month or something like that.

Other times, users indicated that perhaps the tie strength threshold for strong ties was set to low, making for strange inclusions in the Inner Circle.

> [Did any of We Meddle's guesses seem particularly out of place?] Some of them did. [Famous musician] wound up in my Inner Circle. Umm ... [And I suppose you don't know him?] Not so much. I don't think seeing his show counts.

However, by far the most interesting edge cases came from actual relational contexts which fell outside We Meddle's assumptions. For example, the following two quotes showcase two types of relationships that We Meddle misinterpreted.

Some people on Twitter just say stupid things. Or, they might say something that's inaccurate. So there's a few people I see who are not in my Inner Circle, you know, my [real life] group ... people I actually hang out with. But we have had disagreements on Twitter. You know, we argued.

The Inner Circle is actually not super accurate. Yeah, the Inner Circle is basically all the people I used to work with. What's a way better representation of my inner circle is actually the "Flocking Together" group ... Almost all the Inner Circle people are [company] people I used to work with. But I talk with them, sort of irregularly, now that I'm not at [company] anymore. [How close were you when you worked there?] Yeah, I was close to them while I was there, so it's fun to see them here from an old job. Maybe there should like a category for them: like people you used to be close to, but you know, aren't anymore.

Here, we see two very interesting violations of We Meddle's assumptions about relationships online. First, relationships can be intensely *negative*. In Chapter 3, we saw a manifestation of this in an "exgirlfriend" comment. (In fact, I received a some emails from We Meddle users wondering why their expartner appeared in the lists.) The tie strength model does not very accurately capture these kinds of relationships. We can do better here.

The participant's story about the former job showcases how a *biographical break* can color someone's viewpoint. It seems clear that the participant was at one time close to these people, but no longer. We Meddle's computational backend does not handle (or even really generate) sharply discontinuous data. It seems likely that this person felt close his co-workers one month, and then did not the next. (Interestingly, note how we says he says he more closely aligns with another of the lists, a community-detection list. He has moved on.) Future work could do better by trying to resolve these persistent problems. However, while certainly problems, we can actually view persistency as a positive sign for generalization: this model makes the same kind of mistakes the model in Chapter 3 made.

For the sake of completeness, I will also include some users' thoughts on community detection lists, even though it is not the core focus of this dissertation.

The Birds of a Feather [one of the community detection lists], I couldn't figure that one out. However, there's some people that just seemed like they're the more upperechelon-type users. I remember a couple of my friends are on there—just power users, extreme power users.

[Regarding the community detection lists] I'm a pretty active follower of a local sports radio station, and it put those users in there. And I'm pretty sure it got all of them. That was the first list I created in Hootsuite. So that was really cool.

Most of the people I interviewed continue to use their lists to this day. In particular, a number of users that followed lots of people commented that the service particularly helped them.

I'm a pretty active user, you know, I follow over one thousand people ... I think the lists thing is going to be really big. I think you've got something. I think the list thing is going to be huge ... it's going to be huge.

I thought so ... I think it's a really cool idea. The only part I found lacking was the functionality after the lists were auto-generated. [You mean adding?] Yes, adding.

[Have you used or modified the lists you made outside of We Meddle?] I went and deleted one through Twitter. I wanted to drop some of the people so it corresponded to a list I'd made earlier in [another client]. It was almost there.

[Were there many people you wanted to add to the Inner Circle?] Yeah, there were around 20. I saw it as meaning users I actually know. But some of them, I don't want to see their tweets. So it was probably 60% knowing them, 40% content driven. I was dropping people I knew but didn't really care about their updates.

We see an interesting tension, one not explored in We Meddle, about the difference between social relevance and content relevance. We Meddle lives in the social world, and makes almost no attempt to work with content. I think this would be a big win for a revamped We Meddle: a We Meddle that pairs social relevance with content relevance.

For the most part, the users I interviewed did not find much use in the client. It either seemed to weird, or they simply did not want to give up every shortcut they had learned in their existing clients. Also, from my logs it seems apparent that many client users came from Asia, while all of my interview participants came from the U.S. and Europe. However, it seems reasonable to assume that the client may have found some traction in Asia because it's still under the radar of Chinese censors. I wish I could say: no Asian We Meddle users responded to my interview invitations.

[The client] is not really useful for me right now. I keep Tweetie open all day. And, you know, I might be OCD, but I read every tweet. And so for me, something that filters my Twitter feed isn't really valuable because I actually want to see all the content. Maybe for someone who follows more people than they can handle, and they want to see just what their friends are talking about and they don't want to Twitter Lists then maybe. That's what I use the We Meddle lists for.

From discussions like this one, I think the client deserves more UI attention as well as more attention to explaining very clearly what it does for users. Some users, as you see in the upcoming tweets, did fully get the client, but clearly not most. (It also did not help that the We Meddle client required a second click on a small link from the Lists interface. However, I made this decision consciously to collect as much tie strength data as possible.) I'm going to work on revamping the client in these ways, hopefully releasing a new version in the next year (and with a more fanfare and press).

#### Reactions on Twitter and the Web

In addition to the interviews I conducted with We Meddle users, I also include here (Figures 27 through 36) a selection of unprompted reactions from Twitter and the web. They come from a variety of media (e.g., blog posts, tweets, live webcast video), and reinforce themes I have already discussed.



Figure 27. A screenshot from an impromptu reaction to an We Meddle a talk by Prof. Karahalios. The talk was webcast live from the Berkman Center for Internet and Society at Harvard. I captured this portion of the talk and made it available at http://social.cs.uiuc.edu/berkman.we.meddle.mov.



Figure 28. An unprompted reaction from a Twitter user. The fact that We Meddle recreated lists he had already made for himself speaks to computational tie strength's face validity.



Figure 29. Berkman fellow David Weinberger reacts to We Meddle.



Figure 30. A We Meddle user comments on his lists during Prof. Karahalios's Berkman talk.



Figure 31. We Meddle has received many visits from non-English speakers. This tweet roughly translates as: This tool gets your Twitter friends (!?), processes them and afterward gives you lists. When you see the results you can edit however you want.  $^{^{^{^{^{^{^{^{^{^{^{^{^{*}}}}}}}}}}$ 

Iisu con http	ise Ţestoase, ce mi tactele din Twitter p://wemeddle.com	șto filtrează aplicația asta /	*
12:37 A	M Mar 4th via Echofon	n Reply	Retweet
9	OviSirb Ovidiu Sirb		

\_ \_ \_

Figure 32. A Romanian user writes: Oh geez, this application is a great filter for your Twitter contacts.

twitter Home Profile Find People Settings Help Sign o 這是從 http://wemeddle.com/client/ 發 的tweet. wemeddle真不錯, 自動把我的推 友分群, 這樣就不怕重要朋友的偶爾一推
這是從 http://wemeddle.com/client/ 發 的tweet. wemeddle真不錯, 自動把我的推 友分群, 這樣就不怕重要朋友的偶爾一推
的淹沒在一片汪洋裡 about 14 hours ago via We Meddle Retweeted by 1 person
© 2010 Twitter About Us Contact Blog Status Goodies API Business Help Jobs Terms Privacy

Figure 33. A Chinese user writes: This is a tweet from this link [client link]. We Meddle is pretty good, it automatically divides my friends into groups, so that occasional tweet from my important friends won't be buried in an ocean of other tweets.



Figure 34. Another Chinese user writes: This thing is miraculous. I want to know how it works very much. It even classifies my GF's account into Inner Circle, but I never send tweets to her account, and it also classifies accounts I don't care about into outer circle. The 4 categories are very accurate, almost 100%. It is miraculous.

	Cameronfactor w with the MeFites. A any chance?	Home Profile Find People Settings Help Sign out vemeddle.com put you Are you a MeFite by	
	AliMarin Alexandra Marin		
(	© 2010 Twitter About Us Contact Blog S	tatus Goodies API Business Help Jobs Terms Privacy	

Figure 35. We Meddle has started conversations, as well as acting as a way to sort a user's own contacts. In this tweet, a We Meddle user asks the Facebook researcher Cameron Marlow if he participates in the online community MetaFilter. (A "MeFite" is a MetaFilter member.)



Figure 36. In this blog post, a user of We Meddle screencapped the We Meddle interface and commented on it. The user goes on to tell a story about the people in the Inner Circle, even apologizing for those left out.

## CHAPTER 5 INFORMATION DYNAMICS & TIE STRENGTH

[*When is an Edge an Edge?*] Dichotomous distinctions can sometimes be misleading. Many forms of interaction are inherently episodic and occur at variable rates ... This cannot be resolved solely with better data collection or more elaborate statistical techniques. Rather, one must determine whether the relationship under study is sufficiently stable to be well-approximated by a constant function over the period of interest and whether the values taken by this function across pairs are sufficiently constrained to be approximately dichotomous. For relationships known to be highly heterogeneous (such as trade or migration rates), no single threshold may suffice; a weighted graph representation will frequently be more appropriate. More studies that assess the effectiveness of such approximations—and provide concrete, empirically validated guidelines for practice within particular problem domains—would be a welcome addition to the literature. — 2009 *Science* article "Revisiting the Foundations of Network Analysis" (Butts 2009)

Many would argue that the primacy of networks is a core contribution of modern science (Benkler 2006; Borgatti et al. 2009; Jackson 2008; Watts 2004). Networks have been shown to govern many social (Borgatti et al. 2009; Watts 2004), economic (Hidalgo et al. 2007; Jackson 2008), ecological (Proulx et al. 2005) and health phenomena (Cohen et al. 2000; Potterat et al. 2002). A succinct summary: networks are the substrate in which various human processes take shape.

One extensively studied area in the networks literature is diffusion (Bearman et al. 2004; Christakis & Fowler 2007; Fowler & Christakis 2008; Ivkovic & Weisbenner 2007; Klovdahl 2005; Kossinets et al. 2008; Onnela et al. 2006; Rogers 1995; Scherer & Cho 2003). Goods, services, money, diseases, beliefs and information all flow through networks. Almost without fail, these studies focus on the macroscopic properties of networks (e.g., Albert & Barabási 2002; Castellano et al. 2009) and their associated outcomes. Of course, these are powerful and valuable studies. Before the advent of abundant data and the cheap hardware and software to process it, we simply could not do these studies. In this chapter, I hope to make a contribution to in a forgotten area of diffusion: the microsocial, relational processes that govern that make diffusion possible. In other words, here I study how tie strength affects diffusion.

For reasons of tractability, or perhaps due to lack of data and solid constructs, almost all diffusion studies assign relatively simple probability distributions to govern when and where information spreads (Fowler & Christakis 2008; Gruhl et al. 2004; Kempe et al. 2003; Kossinets et al. 2008; Liben-Nowell & Kleinberg 2008). (This includes, and perhaps started with, the popular Susceptible-Infected-Recovered, or SIR, model.) One recent exception stands out. The only diffusion work I am aware of to consider microsocial properties (i.e., tie strength) is the Onnela, et al. (2006) study of mobile phone customers. In this large-dataset study, the authors obtained a large dataset of one month's mobile calls made by a subset of customers in a metro area. They operationalize tie study as the minutes of call time between two people. (This definition seems natural and intuitive to me, given the context of the data.) The paper then *simulates* how information may flow through networks whose edges understand tie strength. This paper is a very valuable point of comparison for my work, and I will return to over the course of this chapter.

Taken together, however, the diffusion literature makes a big simplifying assumption: simple stochastic processes govern node-to-node transmission. How does a disease spread? With constant probability across the network. How does a meme go from one blog to another? Via a power-law rule. Here, I unpack this assumption. If tie strength affects transmission, it's quite possible that standard models vastly skew the macroscopic properties of diffusion, as hinted by (Onnela et al. 2006) and (De Choudhury et al. 2010). In other words, our little operational choices (i.e., what makes a tie) can propagate to have big global consequences at the level of topology and inference.

Most diffusion studies live at what we might call the *global-network* level. They do not mess around with the details of particular kinds of information flowing through particular places in the network under particular conditions. In this chapter, I try to deal with this messiness. With this background in place, I now introduce the three research questions which guide the final piece of my dissertation.

#### RESEARCH QUESTIONS

- R1: Does tie strength modulate the flow of information through a network?
- R2: Does tie strength interact with content as it flows through a network? If so, in what ways?
- **R3**: If tie strength regulates the transmission of information through a network, how does this affect macroscopic properties of diffusion?

Of course, tie strength does not tell the whole story. You may find some web video particularly aweinspiring (Berger & Milkman 2009) and feel compelled to share it with friends. It may have everything to do with the content and your impulse to share it. This will certainly be true in some cases. However, it also seems possible, even likely given findings in the literature, that your relationship with the source affects your likelihood to disseminate content. In Twitter, for instance, diffusion conveys two things: the actual content, plus who wrote it. Not only do people forward information through the network (via the retweet mechanism, explained in detail shortly), but they also push a social signal through the network. If you retweet me, I am aware of it because Twitter makes me aware in its interface. It wants me to see that you've forwarded my content. I argue that it's simply not possible to separate acts of diffusion from the social signals they produce.

Furthermore, the literature also suggests that we get different kinds of information from different sources, and tie strength may tell us something important about this practice. One of the overarching lessons from the early tie strength literature is that tie strength acts differently in different contexts. For
instance, our strong ties are more likely to share our values (Granovetter 1973). Consequently, I hypothesize that value-laden content diffuses differently through a network than value-neutral content (**R2**). Thinking in terms of email, are you more likely to forward the political chain letter that agrees with your political views or the one that does not?

**R3** extrapolates from the immediate, microsocial effects of tie strength on information dynamics to its macro-effects. What effect size do we see? Is tie strength a primary component in the diffusion of information or do we suspect others? Do the effect sizes indicate that tie strength will substantially shape the transmission of information in online contexts? Recent work (Kwak et al. 2010; Sun et al. 2009) has confirmed that the "information cascades" everyone likes to discuss actually happen very rarely. By one estimate (Kwak et al. 2010), only 4% of the time does information travel more than *one hop*. With information staying so close to home, it seems likely that microsocial effects play an even bigger role.

#### Diffusion vs. cascades

In this dissertation, I study *diffusion*, the spread of some thing or concept from one person to another. It is important to differentiate diffusion from a related concept, a *cascade*. Cascades happen when a diffusing process explodes, reaching many people beyond the originator. Think of a how a particular fashion gets popular: it starts in a small community of designers, spreads to influential cities and then goes seemingly everywhere. This is a cascade. Diffusion is a necessary component of cascades, happening at every stage along its spread, but most diffusing processes do not turn into cascades.

Cascades have received lots of attention from the research community because they correspond to things like epidemics, memes and successful viral ad campaigns. I choose to focus on the relatively smaller-scale act of diffusion, the building block of the cascade. From what we learn here, the hope is that we can scale up to events like cascades.

#### METHOD

I study the relationship between information and tie strength using two main pieces of data: retweets and We Meddle's tie strengths. An impediment to a study like this is actually getting your hands on the data. (This is perhaps the reason it has not been done.) It's hard enough to obtain a large network dataset; it's often nearly impossible to get the rich interaction history between the actors linked in that network. However, this is precisely the data I have as a side effect of users coming to We Meddle. In order to render a new interface based on tie strength, We Meddle needs to see tons of interaction data between you and the Twitter users you follow. It looks at many pieces of data (listed in detail in Chapter 4) to make its calculation. But it does not use retweets.

*Retweet* is Twitter-speak for reposting someone else's tweet to your own followers. Figures 37 and 38 show two example retweets at the time of this writing. In many ways, retweets resemble email forwards.

Very often (recent work estimates 52% of the time), the retweet contains a summary or opinion about a link to somewhere on the web. Retweets are the primary way that information diffuses on Twitter.



Figure 37. A retweet using the "built-in" mechanism for reposting tweets. Twitter provides a button (and a formal API mechanism) to redistribute someone else's tweets to your followers. In this case, the user dabbish and five others have retweeted physorg\_com's tweet, which contains a link to the full story.



idsgn Video: Around the world in 80 seconds http://bit.ly /aFa9rJ (via @GOOD) about 3 hours ago via TweetDeck

Figure 38. A retweet using a user-adopted syntax, in this case the *via* syntax. Here, idsgn has retweeted GOOD's tweet about a video. The syntax retweet predates the built-in retweet, and many users still prefer it. This style retains a user's identify: syntax retweets appear to come from the retweeter, not the retweeted, as above.

Figure 39 is a high-level depiction of the method presented here. A crawler visited each We Meddle user's tweets in late April 2010, looking for retweets made in any of three ways: via the built-in Twitter mechanism or via two predominant user-adopted syntaxes. While a complete history of the retweet is beyond the scope of this chapter, I will provide a brief one as background to my work. The retweet first appeared in user syntax form. A Twitter user would encounter someone else's tweet, want to repost it to their own followers, but with no official mechanism a community-standard emerged. "RT @screenname" (and later "via @screenname") arose to denote retweeting. Later, Twitter recognized retweeting as a core practice in its community and built it into their infrastructure, providing a button in their twitter.com interface and a documented API method for accessing it outside of twitter.com. This very brief history sidesteps some community outrage that arose in response to Twitter's move. For now, it is sufficient to say that users mostly use these three methods to retweet. In this study, I look for built-in retweets and the two most dominant user syntaxes, "RT" and "via."

Scanning We Meddle users' tweets this way produced 19,087 retweets with links to web content (i.e., containing the text "http://..."). Here, the link is the thing diffusing. Because users can retweet people they do not actually follow, I cut the dataset to those connected by a following relationship. Since this data collection happened in some cases well after the first We Meddle login, I further constrained the

data to those with non-zero tie strength. This permits very low tie strengths (e.g., 0.0003) but removes a very real confound: the We Meddle user started following this person after they first logged into We Meddle. Restricting the data to those with tie strength greater than zero ensures that we are only considering dyads We Meddle had a chance to evaluate, while only cutting a very small number of ties at 0.

I next forked the data into the three sets I will analyze in this chapter. (Also depicted in Figure 39.) First, I selected all data containing links to known multimedia sites: photo-sharing sites like Twitpic and Flickr, video-sharing sites like YouTube, Qik and yFrog and any direct links to multimedia files, such as ".jpg" and ".mov" files. The multimedia dataset contains 503 retweets (after constraining the following relationship and ensuring nonzero tie strength). After setting these aside, I performed content analysis on the URLs in the remaining retweets, leading to the two other datasets. I chose to analyze web content rather than the tweets themselves simply for robustness: validated tools exist to perform automatic content analysis on web text. Tools for 140 characters are still coming online.

After downloading the referenced content in each retweet, I sent it to Reuters's OpenCalais<sup>23</sup> system for topic classification, an emerging industry standard. Among many services it provides, OpenCalais will produce a list of topics for a page from a predefined list of 17 topics. The OpenCalais topics are: Business & Finance, Disaster & Accident, Education, Entertainment & Culture, Environment, Health, Hospitality & Recreation, Human Interest, Labor, Law & Crime, Politics, Religion & Belief, Social Issues, Sports, Technology & Internet, Weather and War & Conflict. I chose OpenCalais over other systems because it is tuned to web content and incorporates many training sources. Without a hand-labeled test set, it is impossible to confirm the absolute accuracy of OpenCalais; however, many sources report success with OpenCalais. While we will see noise, the hope is that the noise is statistically manageable.

One purpose of this study is to investigate the relationship between content and tie strength as information diffuses through a network. Topics give us one cut at that relationship. Another view is subjectivity. As stated in the research questions, theory would expect subjectivity to vary with tie strength. Topics give us a rough approximation of subjectivity because, for instance, we might expect political and religious news to express subjective views more often than business news. (Topics give us insight into things other than subjectivity, as well.) But to look at subjectivity in greater depth, I used Opinion-Finder (Riloff & Wiebe 2003; Wilson et al. 2005), a system from the University of Pittsburgh, to label subjectivity in the form of subjective sentences and direct subjective expressions. The authors report an accuracy of 74% on a basket of corpora for annotating subjective sentences, and 80% for classifying direct subjective expressions. ("He thought Obama had won the race" is a subjective sentence; "thought" is the direct subjective expression.) After applying all the preprocessing steps above, and only including web content which OpenCalais could label, both the topic-coded and subjectivity-coded da-

<sup>23</sup> http://opencalais.com

# tasets have 4,314 retweets. (The *following-exists* check eliminates approximately 6,000 retweets; inability to annotate topics accounts for the rest.)



Figure 39. A high-level overview of the method for collecting and preparing data for the analysis presented in this chapter. After capturing three types of retweets and preprocessing, I generate the three main datasets I analyze in this chapter: multimedia retweets, topic-coded retweets and subjectivity-coded retweets.

# Retweets as the mark of diffusion

If you're familiar with Twitter, you know that diffusion happens all the time without retweets. Diffusion is built right into the core of Twitter: people network-broadcast information at each other all day long. I may push a link out to everybody who follows me or just write a quick note about my day, and all my followers see it. But they have to watch. If they look away for a moment, or take a phone call, they could miss it. And this is the wrinkle. We have to know that someone read the status update and got the information. If the tweet contains a link (and most do), then we need to know that they clicked on the story and got its meaning.

So, using the retweet as a mark of diffusion is really a method hack. By saying that retweets equal diffusion, we can certainly say that they read the tweet, and almost certainly say that they read the link they forwarded. Other diffusion studies make similar method compromises.

# RESULTS

Retweeting behavior depends on tie strength. In this section, I will use *All We Meddle ties* as the reference distribution because a uniform sampling process among ties (i.e., tie strength plays no role in diffusion) would produce this distribution. Comparing tie strengths among the groups *All We Meddle ties*, *Retweeted ties* and *Multimedia retweeted ties*, Kruskal-Wallis  $X^2 = 207.72$ , 2 d.f.,  $p < 2.2 \times 10^{-16}$ . Post-hoc Wilcoxon rank-sum tests confirm pairwise differences between all three groups: *All-Retweeted W* = 7,222,335,  $p < 2.2 \times 10^{-16}$ ; *All-Multimedia W* = 786,612.5,  $p < 2.2 \times 10^{-16}$ ; *Retweeted-Multimedia W* = 959,161, p = 0.00239. Figures 40 and 41 depict these results in graphical form. We see a skew toward stronger ties among the retweet groups, with multimedia retweeted ties median = 0.422. The reference distribution in Figures 40 and 41 shows a peak around its median. This is substantially lower than the peak we saw in Facebook data in Chapter 3. Two factors explain this: Twitter users often follow many more users than Facebook users have friends, and they have sparser interactions with these users, leading to a distribution with a fatter head.

Both the *Retweeted ties* and the *Multimedia retweeted ties* have three humps: one near 0.2, one near 0.6 and one near 1 (though less pronounced in the case of *Multimedia retweeted ties*). These are at least bimodal distributions, and perhaps trimodal distributions, suggestions different classes of activity. The distributions shown in Figures 40 and 41 have been produced using a Gaussian kernel density estimate, and therefore have some discontinuities removed. Figure 42, on the other hand, shows the probability of retweeting (i.e, diffusion) conditioned on tie strength, and it leaves in jagged parts of the curve. This is meant to contrast with the completely smooth distributions in Figures 40 and 41.



Figure 40. The reference distribution of all We Meddle tie strengths (gray) and the distribution of tie strengths from retweets (filled yellow). If tie strength had no effect on retweeting behavior, the two would closely align. Instead, this graph shows that tie strength affects retweeting behavior, with stronger ties bulging out at near 0.6 and 1. A Kruskal-Wallis test confirms this visual intuition. (Note: these curves represent probability distributions, smoothed with a Gaussian kernel. Each curve is generated from 4,158 data points.)



Figure 41. The reference distribution of all We Meddle tie strengths (gray) and the distribution of tie strengths from retweets containing links to photos, videos and music (filled brown). Retweets with multimedia links look similar to retweets generally, but skews to stronger ties. Post-hoc Wilcoxon tests confirm the difference between media retweets and retweets. (Note: these curves represent probability distributions, smoothed with a Gaussian kernel. Each curve is generated from 4,158 data points.)

The filled portions of Figures 40 and 41 depict the probability of a certain tie strength conditioned on a retweet,  $P(ts \le t | RT)$ , where ts refers to tie strength and RT refers to a retweet. Using Bayes' rule, we can use this quantity to estimate a distribution of perhaps greater interest,  $P(RT | ts \le t) = P(ts \le t |$ RT)  $\cdot$  P(RT)  $\div$  P(ts  $\leq$  t), the probability of a retweet conditioned on tie strength. Figures 42 and 43 show different estimates of the part of this probability, the ratio  $P(ts \le t | RT) / P(ts \le t)$ , essentially the ratio of the area under the filled curve to the area under the unfilled curve in Figures 40 and 41. Figure 42 uses a step size of 0.01 to estimate this ratio, whereas Figure 43 uses a step size of 0.02. That is, Figure 42 moves in increments of 0.01 down the tie strength axis to estimate the ratio using counts from the dataset, where Figure 43 moves in 0.02-sized increments. Both figures show complex distributions with knots where the data substantially changes shape. Even in the smoother Figure 43, a Multivariate Additive Regression Splines (MARS) fit of the data finds three separate segments of the ratio, each with a different slope. MARS fits a piecewise linear approximation, and gets progressively better as tie strength grows: overall  $R^2 = 0.834$ , first segment  $R^2 = 0.538$ , second segment  $R^2 = 0.932$ , third segment  $R^2 = 0.887$ . I did not want to overfit these data, but it seems that future work might find traction modeling the segment between 0 and 0.2 logarithmically. In the bumpier Figure 42, on the other hand, we see interesting nonlinearities at the endpoints, suggesting different effects for very weak ties (very close to zero) and very strong ties (very close to one). These jumps at the endpoints, plus the MARS fit in Figure 43 lead me to introduce the five groups we will consider later in this chapter: very weak ties, weak ties, medium ties, strong ties and very strong ties.



Figure 42. The *unscaled* pseudo-CDF of retweet probability conditioned on tie strength. This curve uses Bayes's Rule to turn the two figures above into  $P(RT | ts \le t)$ . This graph is essentially the ratio of the filled graph in Figure 40 to the gray line. It is unscaled because this curve leaves out P(RT), the baseline probability of retweet. However, as a constant (e.g., it's median across users) probably suffices for P(RT), this curve will drive overall estimates of P  $(RT | ts \le t)$ .

The curve illustrates how extremely weak ties get retweeted more often than expected. (The smoothing process in the previous figures obliterated it.) With the exception of extremely weak ties, weak ties generally suffer in retweets. This curve shows lots of detail due to its 0.01 step size. However, we see a somewhat steady rise between 0.2 and 0.6, with a jump near 0.6 and a slower rise afterward. Extremely strong ties (e.g., near 1) get a big bump too.



tie strength

Figure 43. Another estimate of the unscaled pseudo-CDF of retweet probability conditioned on tie strength, this time with a larger step size. The step size in this figure is 0.02, whereas the previous figure uses 0.01. Notice how the effect for extremely weak ties goes away, as that one data point  $[P(RT | ts \le 0.01)]$  has been smoothed in a larger bin. The overall trends remain the same. This curve provides more robust estimates, but also sacrifices detail and interesting discontinuities.

The red line shows a Multivariate Additive Regression Splines (MARS) fit of the data. In this case, MARS learns a simplified piecewise linear approximation of the curve. It segments the data into three parts, each with distinct slopes: segments I will call *weak ties, medium ties* and *strong ties*. Based on the high variability we see at the enpoints of Figure 42, we might add the segments *very weak ties* and *very strong ties*.



Figure 44. The method for estimating P(RT) from 635 We Meddle users. Over a one week period, I obtained the size of each user's tweet stream, the number of tweets that would have come through their client had they been logged in the entire time. Counting the number of retweeting relative to the overall size of the stream gives you an estimate of P(RT). It's not the whole story, however, since users may be logged out or simply not paying attention. It's debatable whether this should factored in or not, as I have done here.

Estimating P(RT), the missing part of the equation, is a tricky business. While the ratio P(ts  $\leq$  t | RT) / P(ts  $\leq$  t) will play the largest role in driving P(RT | ts  $\leq$  t), having P(RT) will give us a complete picture. To estimate it, you have to know two things: how many tweets streamed past a user's eyes and how many they subsequently chose to retweet. The latter is easy; the former is hard.

The approach I take here is to estimate the size of a user's tweet stream (the stream they would see from their perspective as they view Twitter) to calculate P(RT). I depict this process graphically in Figure 44. In particular, I estimate the size of the tweet stream for 635 We Meddle users for the week between April 26 and May 3, 2010. (The choice of one week is a practical constraint, the length of time Twitter's search API keeps tweets.) Hitting every user the 635 We Meddle users follow would have required a complex distributed crawl of over 1.5 million API requests. Instead, for each We Meddle user, I gather the number of tweets made by a random 150-user subset of their network over this one week period and inflate the size of their tweet stream appropriately. (This sampling process saves about 1 million, probably unnecessary, API requests.) In cases where the user follows less than 150 users, I visit them all. Figure 45 shows the distribution of these users' individual P(RT) values, with a median of 0.0001262704 (shown to many digits to help other researchers). However, I suggest using a sum total

measure, that collapses activity from each user into a single number. Using this method, P(RT) = 0.0001426229. Scaling any value in Figures 42 and 43 by this number gives  $P(RT | ts \le t)$ .



Figure 45. The probability distribution of P(RT) at user level, from 635 We Meddle users. The median of this distribution is 0.0001262704, but I collapse all activity into one estimate of P(RT) = 0.0001426229.

# *Tie strength and content*

Having examined retweeting behavior in the aggregate, I now turn to how tie strength affects the way different kinds of content diffuses. In this subsection, I look at tie strength's impact on topic, as measured by Reuters's OpenCalais topic classification system. In the second half, I study how tie strength interacts with subjectivity expressed in the text. Table 12 shows the results of 17 logistic regressions, one for each of the 17 OpenCalais topics listed earlier. In each, tie strength is a predictor along with the following controls: usage (log (RTer's statuses count) and log (RTed's statuses count)), the retweeter's baseline rate of retweeting (as a percentage of their status count), fixed effects for repeated individuals in the sample, day of the week, and day or night. Each control represents a guess of mine at an alternative hypothesis: for example, perhaps business news predominantly diffuses on Monday (day of week effect), after a weekends where comparatively little happens.

Since I performed 17 simultaneous tests, I controlled the family-wise error rate with a Bonferonni correction, making  $\alpha = 0.05/17 = 0.00294$ . With a dataset from the internet, we need to worry about artificially low p-values. In other words, take any effect size, supply a big enough *N*, and you have significance. So, I will concentrate on effect sizes here. However, note that this dataset is certainly not huge by internet standards (*N* = 4,158). Also, the tie strengths and the dependent variables contain measurement error.

Tie strength has a non-random, negative effect on political pages during diffusion, b = -0.372, z = -3.45, p = 0.00056. The negative effect here surprises me. As it seems natural to link political web content with value-laden content, I expected a positive effect. I revisit this when I examine subjectivity, without needing a proxy like topic. I offer the following alternate explanations for tie strength's negative effect on the diffusion of political pages: perhaps we are really measuring "breaking news" and not politics as OpenCalais thinks; or, perhaps political news circulates particularly tightly in networks of strong ties and people only need to forward political content from weak ties.

Tie strength has weaker, but potentially random effects on *Business & Finance* (negative, b = -0.101, p = 0.0626), *Entertainment & Culture* (positive, b = 0.0982, p = 0.06518) and *Technology & Internet* (negative, b = -0.0938, p = 0.0519). These effects all seem to line up with existing theory: Business and Technology pages probably represent value-neutral content, whereas Entertainment and Culture pages clearly have value and lifestyle embedded within them. Compared to tie strength's effect on *Politics*, these effects are substantially weaker. As far as I know, this is the first work to examine tie strength's effect on content.

Business & Finance Topic	b	Std. Error	Z	р	LRT	р
log(RTer's statuses count)	1.92	29.30	0.066	0.948	0.00	0.947
log (RTed's statuses count)	-0.037	0.056	-0.665	0.506	0.40	0.506
RTer's baseline RT rate	-65.1	70.1	-0.929	0.353	1.10	0.284
Individual effects					583.90	< 0.001
Day of week					13.70	0.005
Day or night	0.143	0.111	1.292	0.196	1.70	0.195
Tie strength	-0.101	0.054	-1.862	0.063	3.50	0.061

**Disaster & Accident Topic: whole model rejected** 

Education Topic: whole model rejected

#### **Entertainment & Culture Topic**

log (RTer's statuses count)       -29.120       18.640       -1.562       0.118       2.40         log (RTer's statuses count)       0.021       0.057       0.365       0.715       0.10	0.119 0.715
$1 \circ \pi (DT \circ d' \circ totulo \circ \circ o o t ) = 0.021 = 0.057 = 0.265 = 0.715 = 0.10$	0.715
$\log(\text{Red S statuses count})$ 0.021 0.037 0.305 0.715 0.10	
RTer's baseline RT rate         49.370         38.280         1.290         0.197         1.50	0.216
Individual effects 796.50	< 0.001
Day of week 11.00	0.090
Day or night         0.051         0.108         0.478         0.633         0.20	0.633
Tie strength <sup>†</sup> 0.098         0.053         1.844         0.065         3.40	0.066

Environment Topic: whole model rejected

Health Topic: whole model rejected

# Hospitality & Recreation Topic: whole model rejected

#### Human Interest Topic

log(RTer's statuses count)	85.050	46.440	1.831	0.851	4.90	0.028
log(RTed's statuses count)	0.013	0.063	0.210	0.834	0.04	0.834
RTer's baseline RT rate	201.400	81.770	2.462	0.014	6.10	0.014
Individual effects					634.20	< 0.001
Day of week					8.00	0.237
Day or night	0.056	0.124	0.450	0.653	0.20	0.653
Tie strength	0.093	0.063	1.469	0.142	2.10	0.143

# Labor Topic: whole model rejected

Law & Crime Topic						
log (RTer's statuses count)	0.028	0.105	0.265	0.791	0.07	0.791
log (RTed's statuses count)	0.150	0.098	1.541	0.123	2.43	0.119
RTer's baseline RT rate	-0.829	0.212	-3.912	< 0.001	24.33	< 0.001
Individual effects						
Day of week					11.18	0.083
Day or night	-0.715	0.182	-3.922	< 0.001	15.44	< 0.001
Tie strength	-0.165	0.098	-1.678	0.093	2.92	0.088
Politics Topic						
log (RTer's statuses count)	-93.070	107.800	-0.864	0.388	0.69	0.405
log (RTed's statuses count)	0.011	0.112	0.099	0.921	0.01	0.921
RTer's baseline RT rate	36.150	224.700	0.161	0.872	0.03	0.874
Individual effects					639.85	< 0.001
Day of week					11.04	0.087
Day or night	-0.059	0.202	-0.293	0.770	0.09	0.770
Tie strength	-0.372	0.108	-3.450	0.00056	12.71	0.00036

# Religion & Belief Topic: whole model rejected

Social Issues Topic: whole model rejected

Sports Topic: whole model rejected

# Technology & Internet Topic

log (RTer's statuses count)	-3.655	24.640	-0.148	0.882	0.02	0.881
log(RTed's statuses count)	0.096	0.050	1.926	0.054	3.700	0.054
RTer's baseline RT rate	0.129	50.430	0.003	0.998	< 0.001	0.998
Individual effects					1035.30	< 0.001

Day of week					6.80	0.341
Day or night	0.005	0.095	0.048	0.961	< 0.001	0.961
Tie strength	-0.094	0.048	-1.944	0.052	3.80	0.052

Weather Topic: whole model rejected

War Topic: whole model rejected

Table 12. The results of logistic regressions to predict automatically-inferred topics of retweeted URLs. Each regression controls for the same things, including fixed effects (modeled as dummy variables for each screen name). *LRT* refers to the  $X^2$  liklihood-ratio statistic. After controlling the family-wise error rate with a Bonferonni correction for the 17 topics ( $\alpha = 0.05/17 = 0.00294$ ), I can only confirm an effect (negative) of tie strength on *Politics*. However, tie strength also seems closely related to the *Entertainment & Culture* (positive) and *Technology & Internet* (negative) topics. Models where overall variance did not improve significantly are marked as "whole model rejected."

My arguments around topics concern values: politics stories would seem to have them, whereas business stories would seem to leave them out. However, we see a negative effect of tie strength on *Politics*. To understand this better, I looked directly at *subjectivity* in the text, as automatically labeled by the OpinionFinder system. To recap, OpinionFinder labels subjectivity in two ways: subjective sentences and direct subjective expressions. I measure tie strength's effect on both. Here, I use the same basket of controlling variables, but add a control for the sensitivity of the classifier (either raw bytes or objective sentences, to control for the page's length). I apply negative binomial regression, to model a count. Table 13 shows the results of these regressions. They are strange, to say the least. Tie strength exerts a negative effect on subjective sentences (b = -0.0389, z = -2.67, p = 0.0077), but it exerts a positive effect on direct subjective expressions (b = 0.02, z = 2.065, p = 0.0389). They are, however, small effects. I have included both results for the whole dataset and a restricted subset (e.g., subjective sentences greater than 5) to show how tie strength's effect becomes apparent when we cut out the noisy small documents at the bottom of the range (e.g., automatically generated error pages for expired URLs).

The conflicting directions of tie strength's effect perplexed me. Since you have to be both a *sentence* and *subjective* to be a subjective sentence, perhaps we are seeing stronger ties forwarding pages with more informal language. In other words, stronger ties forward informal language that confuses the classifier. At present, I can only call these results inconclusive. Perhaps tie strength is interacting with topics or some other unmeasured aspect of language.

Subjective sentences	b	Std. Error	Z	р	LRT	р
log(RTer's statuses count)	-5.23	6.56	-0.797	0.425	0.70	0.394
log (RTed's statuses count)	0.034	0.017	1.973	0.049	3.70	0.055
RTer's baseline RT rate	17.0	19.4	0.873	0.383	0.60	0.437
Individual effects					611.40	< 0.001
Day of week					13.40	0.037

Day or night	-0.032	0.033	-0.986	0.324	0.90	0.336
log (objective sentences)	0.787	0.017	46.706	< 0.001	2715.50	< 0.001
Tie strength	-0.028	0.016	-1.700	0.089	2.70	0.1
Subjective sentences > 5	b	Std. Error	z	р	LRT	р
log (RTer's statuses count)	-5.59	5.80	-0.963	0.336	1.10	0.298
log (RTed's statuses count)	0.008	0.015	0.522	0.601	0.30	0.609
RTer's baseline RT rate	31.9	20.9	1.525	0.127	2.40	0.124
Individual effects					535.90	< 0.001
Day of week					15.00	0.020
Day or night	-0.025	0.029	-0.888	0.375	0.80	0.384
log(objective sentences)	0.743	0.015	50.342	< 0.001	3094.50	< 0.001
Tie strength	-0.039	0.015	-2.665	0.008	6.70	0.009
Direct subjective expressions	b	Std. Error	z	р	LRT	р
log (RTer's statuses count)	0.40	3.98	0.099	0.921	0.01	0.925
log(RTed's statuses count)	0.036	0.010	3.543	< 0.001	12.00	< 0.001
RTer's baseline RT rate	14.1	7.4	1.899	0.058	4.00	0.059
Individual effects					774.00	< 0.001
Day of week					7.00	0.360
Day or night	-0.003	0.019	-0.170	0.865	0.03	0.866
log(bytes)	1.403	0.011	128.684	< 0.001	22022.00	< 0.001
Tie strength	0.003	0.010	-2.665	0.781	0.08	0.783
Direct subjective expressions >	b	Std. Error	z	р	LRT	р
log (RTer's statuses count)	1.99	3.66	0.544	0.587	0.30	0.606
log(RTed's statuses count)	0.008	0.015	0.522	0.601	1.00	0.323
RTer's baseline RT rate	13.0	6.8	1.897	0.058	3.40	0.066
Individual effects					515.10	< 0.001
Day of week					14.40	0.025
Day or night	-0.025	0.018	-1.378	0.168	1.90	0.171
log(bytes)	1.225	0.015	83.539	< 0.001	8440.40	< 0.001
Tie strength	0.020	0.010	2.065	0.039	4.20	0.041

Table 13. The results of negative binomial regression on estimates of subjective language in retweeted URLs. I used the Opinion Finder system to automatically label both subjective sentences and direct subjective expressions in the text of the retweeted URLs. (Both terms are defined in the main text.) The regressions control for the same basket of variables as in the topic regressions, plus controls the sensitivity of the classifier: logged objective sentences and logged bytes. Without filtering the data, tie strength exerts no influence on subjectivity. However, filtering out classifier noise on small documents does yield effects, though not strong ones and in opposite directions. One interpretation is that tie strength positively affects subjectivity but negatively affects automatically-labeled sentences—meaning that it positively affects the amount of informal language.

#### DISCUSSION

These results show how real-life diffusion depends on tie strength. Retweets have a median tie strength 40% higher than what we would expect at random, and multimedia retweets have a 64% greater median tie strength. These are big differences. In (Onnela et al. 2006), the authors illustrate how diffusion processes stall in tight-knit communities if diffusion is linked to tie strength. In their simulated work, they link diffusion to a smooth estimate of call time between two people. My results, however, show that the picture is not nearly so simple. The distribution is not smooth, exhibiting a bimodal quality and perhaps a trimodal quality. While certainly understanding tie strength helps you understand diffusion, modeling it as direct variation completely misses the nonlinearity.

The probability of diffusion conditioned on tie strength, as a corollary, exhibits this same jaggedness. We see rather unexpected jumps at the endpoints, people we probably follow strictly for information (e.g., very weak ties) and people we follow because we are really close (e.g., very strong ties). In between these two extremes, we see different effects at different tie strength ranges. Future work modeling diffusion might find quite a bit of traction with this simplifying assumption, in particular the MARS fit of the conditional probability. Researchers are welcome to embrace the complexity of the very bumpy graphs in Figures 42 and 43, but the simplicity of the MARS fit is probably more usable and general.

Tie strength also interacts with content, evidenced by tie strength's effect on the diffusion of political pages and on subjectivity. As far as I know, this is the first study to examine content (e.g., topic, subjectivity and multimedia) as it diffuses. However, the particular directions of tie strength's influence surprised me. From the literature, I suspected that political pages would diffuse through strong ties. The exact opposite happened. I hypothesized that tie strength would predict higher subjectivity, but we see surprising results: tie strength has a downward effect on subjective sentences and an upward effect on subjective expressions. The noisiness of computational tie strength, topic classification and subjectivity could contribute to these results. Perhaps larger datasets and different content measurements could tease this apart.

In one respect, this is a purely quantitative analysis of data. But how might we leverage it when building systems? Looking at Figure 42, it seems that we might predict greater levels of interest in links from very weak ties and links from very strong ties (as evidenced by users' tendency to retweet them.) For

instance, if a user's been logged out and has missed many tweets, perhaps finding tweets from links from very weak ties and very strong ties makes sense, and gives them an opportunity to forward it on.

## Note on the Onnela et al. graph

Onnela et al. produced a graph showing how different a diffusion process looks when modulated by tie strength, reproduced in this dissertation as Figure 10. I considered replicating it here to show how tie strength *actually* affects diffusion processes. However, the work presented in this chapter paints a much more complex picture than I first expected. Yes, tie strength is predictive of diffusion, sometimes strongly. However, it does not explain everything. The Onnela graph, while instructive, suggests that tie strength alone predicts diffusion; the reality appears much more complex. This chapter has shown to what exact extent tie strength affects diffusion in the real world. What it does not show, however, is what other factors fill in the gaps.

# CHAPTER 6 LIMITATIONS, FUTURE WORK & CONCLUSIONS

Where does computational tie strength break down? How far can we stretch current interface models to support relationships? What else can we learn from computing tie strength? In this chapter, I ask these questions, and examine the limitations of what I have presented here. I hope that this process sheds light on where future work should go next. I also conclude the dissertation in this chapter, recapping what we've learned about computing tie strength.

#### MODELS AND PREDICTORS OF TIE STRENGTH

There are many possible ways to model tie strength. What's the best way to do it? What are the best predictors? What's the minimal set (both in raw number and in ease of collection)? For instance, the model I present in Chapter 3 has a fairly impoverished view of time. We could probably do better. A couple of highly important predictors, the duration and recency of a relationship, do move with time. But, most predictors do not. How much does tie strength change from day to day, month to month, or year to year? These are important questions, and we need more study to answer them. From a systems perspective, the answer is simple: recompute it. We Meddle handled it this way. But how often should we do it? Projecting ahead, we might consider explicitly modeling time decay with a decay function, we discount future returns in financial projections.

One point is clear, however: this particular model, the one presented in Chapter 3, yields relatively stable temporal data. Before committing to tie strength's role in diffusion (Chapter 5), I flirted with analyzing how and why tie strength changes over time, particularly focusing on discontinuities. For example, can we infer a biographical break from discontinuities in tie strength, such as break-ups and new jobs? But, it turns out I have almost no discontinuities in my data. (I estimated change over time by having We Meddle spit out tie strength at monthly intervals, instead of calculating it only for "now.") Other models might generate the kinds of discontinuities we expect from real life relationships, but this one does not. We should explore these issues more deeply.

As I discussed briefly in Chapter 3, I can envision other profitable (and easily gathered) predictors. For example, we might look at what kind of language someone uses to address another person, e.g., "Dear Dr. Smith." Or we might consider how long it takes to respond to someone. Intuition says that very quick replies signal a status imbalance. For example, I usually get back to my committee and advisor very quickly. These two features require little collection effort, but I suspect they will lead to big payoffs. Practically speaking, I had no way to include new variables in We Meddle's model: my coefficients would have been blind guesses. I had no way to access who-friended-who in Facebook. This still troubles me. Think about the rich social information contained in a face-to-face handshake: who offers first, the look on their face, how long the other person makes the initiator wait, etc. For the same reasons, I think we'll see rich information fall out from this subtle signal.

And, systems continually change. Facebook has changed (some would say substantially) since my 2008 study in Chapter 3. At the interface layer, Facebook has added the ability to comment directly on statuses and other media, such as photos and videos. In 2008, users did not have this option; every interaction took place in a threadless environment on the Wall. It seems from my personal experience that practices have changed in response (and history suggests that this would almost certainly be true). The question is: How much? The model seemed to work for We Meddle, built a year and a half after the Chapter 3 model. I think this is a particularly deep question for future work: To what extent are online communities immune from design changes? In other words, how resilient are our findings to design changes?

Although I take the first step on generalizability, it's only the first step. Many other social media remain. It's an open question as to how or even if tie strength can be reconstructed in all of them. Also, as I said earlier, I scoped this dissertation away from recommender systems, but it seems like a perfect fit.

Finally, it's hard to see this work and not think about privacy. It's as easy to build a better system for users with tie strength as it is to apply it toward viral marketing. At the moment, I don't have a good answer. We leave a rich digital trail online. Users see enough value in these tools to put a large part of their social lives in them. (Do users realize how much? Probably not.)

# Alternate models of tie strength

I wrote in some length in Chapter 3 about other ways we might conceptually model tie strength. What way is best? In this dissertation, I do not prove that my model most accurately reflects tie strength. I leave that to other work. Here, however, I would like to discuss specific, different machinery that could improve model accuracy and robustness.

From the analysis of errors in Chapter 4, the network component of the model looks like a fine target. Future work might consider first aggressively decomposing the network into communities before layering tie strengths. Second, I suspect that modeling cross-ego-network ties may yield that most bang for the buck, given the findings we see in Chapter 4. That is, we could probably model better if we had interaction data between alters as well (e.g, between your officemate and your boss). Of course, obtaining this data and working with it in real-time remains a challenge.

However, we could envision even more radical departures from the model presented in Chapter 3. A purely computational approach that cares nothing for generalizability or domain adaption might try an

SVM regression model: let it learn the complex, nonlinear interactions between predictors. My experience suggests that you will get highly accurate results. However, I think that's the wrong direction. As I stated previously, I think feature selections matters most. But also, we could rethink how we formulate tie strength itself. I recently found myself reconsidering the basic idea that tie strength lies somewhere between 0 and 1. At first, it seemed quite natural. Yet, you do not have to look long for a counterexample in your own life: just think of someone who you interact with only out of necessity. You may not like them at all. The model presented in Chapter 3 does not accommodate relationships like this, something the literature calls *signed networks*. In other words, in these networks ties can be negative. Most work in signed networks models clearly negative relationships like voting *against* someone in an election. That counts as a negative tie. Again, the model in Chapter 3 cannot accommodate negative ties, and this bears on the network component of the model. Tie strengths only get bigger because of the people bound together in an ego network. What if they could also get smaller, because of negative tie strengths? I find this fascinating. Would you gain something by modeling negative ties? How many fall onto the negative side? Would we find different proportions in the negative range if we compared online and offline relationships?

#### RENDERING SOCIAL MEDIA SOCIALLY

In Chapter 4, I present a re-rendering of Twitter by its underlying social relationships. But how far can we push it? I explored only a few points in the design space. In an early iteration, I found that you could not push the social rendering to the point that it obliterated time. Users rejected it.

For instance, what if we held email from non-work, weak ties until the end of the workday? (Of course, you would leave it to users to adopt this into their own practices, instead of forcing it on them.) This is an example of how we could build tie strength directly into the medium—not just the interface. We Meddle shows that tie strength can help real people use social media better. But there are still so many ways to explore tie strength in rendering social media. Next, I present some short scenarios, each featuring tie strength as a main design element.

#### 1. Prioritize notifications by the user's relationship with the sender of the message.

Over the years, many researchers have attacked the *notification problem:* How and when do we give users new information (Iqbal 2008)? Sometimes new information comes from the system itself (e.g., the operating system notifies you that you will run out of battery soon), but probably more often it comes from other people. Think of email, Twitter, instant messages, phone calls, etc. Most work on notification management, however, frames it as a user-modeling problem: How can we best understand the user's task and mental state to determine the right time to deliver the message? Of course, that's part of the equation. But, I argue that we need to consider the social context of notifications to build robust interruption models. What's the user's past

history with the person who wants to interrupt? Think about executive assistants in real life. Do you think they hesitate to put the CEO through even if the boss is busy?

So, consider the following scenario. I am working on a report, but I am also logged into an email client programmed to check my mail every ten minutes. When it finds new mail, it alerts me with a notification bubble, disrupting me. (Most current email clients, including web-based clients, work this way.) On the one hand, this is a user-modeling decision: Am I currently absorbed in writing, seemingly non-interruptible? On the other hand, this is a social modeling problem: Who sent the email and what is it about? Imagine that before I started writing, I specified an *only-important-people-get-through-to-me* policy because I really need to write. (This is not outlandish; well-known tools already let users turn off the whole internet so they can get work done<sup>24</sup>.) As emails arrive they filter through this policy, backed up by tie strength; only the high tie strengths get through.

# 2. Find new friends.

Often in modern social media, users articulate their connections to one another: friends in Facebook, followees in Twitter. Setting aside the startup problem for a moment (e.g., "How do new users find their initial set of friends?"), how can you suggest new people for your users to befriend? Where is the right place to look? The most popular people? The people with highest mutual connections, but not already connected directly? The prettiest people?

All these strategies make some amount of sense. Good strategies will almost certainly use a mix, but I think tie strength could play a strong role. For example, consider using theories from structural balance (Heider 1946) alongside computational tie strength. Imagine a triad in a social network where two pairs (but not the third) have connected and formed strong relationships, ones the system has inferred by computing tie strength. Figure 46 illustrates this idea.



Figur 46. A illustrative triad clipped from a broader social network where one node is strongly tied to two others, with tie strengths 0.92 and 0.87. Ideas from structural balance indicate that suggesting a completed triad is probably a bad idea.

We might consider suggesting that the two unconnected people form a friendship. They share a

<sup>24 &</sup>lt;u>http://macfreedom.com</u>

common friend after all, a common best friend. However, the lessons from structural balance suggest that this is probably a bad idea. To see why, consider that these two people probably know each other, or at least know about each other. Structural balance suggests that this triad should not exist very often: it causes too much dissonance for the node in the middle. Perhaps there is a reason the gap has persisted over time. They seem to have had ample opportunity to form the friendship and dissolve the dissonance. But they have not. Perhaps it's an implicitly negative link disguised as a gap.

#### 3. Bubble "oldies but goodies" to the top.

As much as social media companies may want us to live our lives staring at their sites, we do not. We go elsewhere on the internet and even occasionally leave our machines. When we come back, all time-based social media just show a linear list of everything that has happened since we left, going on for pages. The longer our departure, the less likely we read it all. But, it does not have to be this way.

A site might find the one or two "oldies but goodies" and move them to the top to ensure that we see them. You can see a place for tie strength here. Perhaps the algorithm looks for an emotionally charged message from a strong tie. Or, perhaps the algorithm selects a highly-commented link from a weak tie. You can see use cases for both. I envision these living in a little floating window for a short time, giving users an opportunity to see it or say they don't care. The system could then learn from these corrections.

These represent only three design points in this space. They illustrate the basic concept: we can use computational tie strength lots of places to potentially make social media better. I have only scratched the surface with We Meddle, but I hope it illustrates how tie strength can change social media designs.

#### Generalizability

All these ideas hinge on generalizability. Do they transfer to other sites under different circumstances? In Chapter 4, I show that this computational tie strength model generalizes past the single site for where it trained.

We need to know more about where computational tie strength generalizes. Does it work in email? Does it work in IM? Perhaps it will need modifications. What modifications? Although clever people will certainly come up with other solutions, future work may consider the approach I took in We Meddle: offering a baseline generic model that users tune via feedback.

#### LEARNING ABOUT OTHER THINGS WITH TIE STRENGTH

In this dissertation, I showed that tie strength affects the diffusion of links across the Twitter network. But many questions remain, some concerning my approach in this dissertation and other, new questions that tie strength could someday tackle. For instance, my data are simply too small to say much about cascades. Most diffusion, 96%, is just a hop between two people. I would like to apply tie strength to the study of cascades, but simply cannot do so with the data I have. If only 4% of retweets take a second hop, with even fewer going farther, I would need many more than the 20,000 retweets by We Meddle users to show anything of real substance. I hope that future researchers with bigger datasets do it.

What other things could tie strength help us study? It seems natural to ask if a certain mix of ties creates specific effects on sites. For example, what mix of ties keeps users on sites the longest? Mostly strong? A core of strong, but otherwise mostly weak? No effect? You can vary the question to incorporate time. Do users go through phases as their tie makeup fluctuates in time? How? Why? All these questions could have profound effects for the way we see online communities and how commercial operators evaluate them. Or, instead of asking questions about survivorship in online communities, we might ask questions about health and well-being. We could re-ask the same questions as Fowler & Christakis (2008): Does emotion flow through tie strength in online networks? Do certain mixes of ties promote prosocial behavior online? Or activity? Or production? It seems that tie strength could play a role in all these studies. Chapter 5 only scratches the surface.

#### CONCLUSIONS

In this dissertation, I have documented a specific mechanism by which tie strength manifests itself in social media. I believe this work addresses fundamental challenges for understanding users of socio-technical systems. How do users relate to one another in these spaces? Do the data left behind tell a consistent story, a story from which we can infer something meaningful? Can we build something with it? This dissertation shows three concrete things. We can infer tie strength from social media. We can use it to dramatically rethink social media design. We can use tie strength as a tool to learn about other things we care about, like diffusion. To conclude this dissertation, I'd like to revisit each of the findings, outlining some of the major results and outstanding questions.

#### 1. We can infer tie strength from digital traces in social media.

The model presented in Chapter 3 incorporates over 70 carefully-chosen, theoreticallymeaningful indicators of tie strength found in a popular social media site, Facebook. It performs with high accuracy, and it informs the tie strength literature by telling us which things matter in which quantities (e.g., Does *duration* matter more or less than *intensity*? By how much?) Due to its relative simplicity, I guessed that I could port it to another online community. I built an application called We Meddle that maps computational tie strength onto Twitter. The findings from Chapter 4 show that this naive adaptation of the Facebook model to a new community generalizes: its errors match the errors in the Facebook model. Chapter 4 also studies the model's mistakes in terms of its input predictors, a useful thing for future models. And yet questions remain. In Chapter 3, I show how interviews turned up instances of what I termed "asymmetric friendships"—friendships dramatically skewed toward one person. How can we fix this modeling problem? It remains an unanswered and attractive target for future research. I propose one potential solution: politeness. The allure of politeness is how little data you need to observe if you know what to look for: you only need to see textual data from one side of the relationship to infer asymmetry. I hope to see this line of work developed.

#### 2. We can use tie strength to make social interfaces better.

Chapter 4 presents We Meddle, an open application for Twitter users. It applies the tie strength model presented in Chapter 3 to a user's contacts and interaction history in Twitter, and is the first application I am aware of to put a calibrated relational model at the heart of its design. Along with the interactive quantitative feedback from users that tells us the error bounds of the model, qualitative feedback suggests that many users find real value with it. Over 1,300 people from around the world have used We Meddle with no coercion or payment; they used it because they thought they would find value in it.

But we can find other applications for tie strength in social interfaces. I discuss some of them in this chapter. Perhaps we do not have to think so linearly; we could bubble old, but salient messages to the top of a linear interface. I like approaches like this one because they meet users halfway—they do not destroy their linear mental models, but they open up news kinds of interactions with social media. But, many other interesting questions remain. How can we best expose tie strength as a controllable tool? How do express it in terms users understand?

#### 3. We can use tie strength as a tool to learn about other things we care about, like diffusion.

In Chapter 5, I presented work showing the diffusion is a function of tie strength. That is, tie strength has predictive capacity about where we will see diffusion. Diffusion studies almost never take relational strength into account. In Chapter 5, I unpacked this assumption. This is the first work I am aware to show that tie strength affects real-life diffusion practices. Moreover, tie strength seems to interact with political content and with subjective content as it diffuses. These findings suggest very different macroscopic properties than what the literature suggests, and they hint at a nest of complexity below the surface. With more studies building on the findings presented here, we may come to better understand that complexity.

I think we can learn about many other things using tie strength. What if we look at contributions to online sites in terms of the distribution of ties? Or well-being as a function of ties in online health communities?

# APPENDIX A LIST OF PREDICTIVE VARIABLES

A list of the raw predictive variables used in Chapter 3's model, categorized and annotated with the author primarily responsible for introducing it:

Amount of time (Granovetter)	Short description	Facebook specificity
first communication	proxy for friendship creation	not Facebook specific
Emotional Intensity (Granovetter)		
number of pokes	low overhead comm.	somewhat Facebook specific
number of status updates	informal, broadcast comm.	not Facebook specific
friend's number of status updates	informal, broadcast comm.	not Facebook specific
number of inbox messages exchanged	higher overhead	not Facebook specific
wall message length	proxy for depth conversation	not Facebook specific
inbox message thread depth	see whittaker, terveen,	not Facebook specific
number of inbound wall posts	frequency of public comm.	not Facebook specific
number of outbound wall posts	frequency of public comm.	not Facebook specific
number of inbound picture comments	times friend commented	not Facebook specific
communication recency	in any channel	not Facebook specific
number of outbound picture comments	included in comm. recency	not Facebook specific
number of inbound tagged in note	included in comm. recency	somewhat Facebook specific
number of outbound tagged in note	included in comm. recency	somewhat Facebook specific
Intimacy (Granovetter)		

inbox message words signaling intimacy	aggregate of all below	not Facebook specific
» family words	brother, sister, mother	not Facebook specific
» swear words	asses, asshole, bastard	not Facebook specific
» friend words	buddy, colleague, companion	not Facebook specific
» sexual words	homosexual, horny, hug	not Facebook specific
» work words	faculty, fail, fax	not Facebook specific
» leisure words	birdie, blackjack, blockbuster	not Facebook specific
» home words	neighbor, oven, patio	not Facebook specific
» money words	dinero, discount, dividend	not Facebook specific
» body words	muscular, naked, nasal	not Facebook specific
» religious words	yiddish, zen, zion	not Facebook specific
» health words	sickness, sinus ,sore	not Facebook specific
» sum of all minus work words	work opposite of intimacy?	not Facebook specific
participant's number of friends	articulated friends	not Facebook specific

friend's number of friends	articulated friends	not Facebook specific
number of gifts from friend	small token of appreciation	Facebook specific
number of gifts to friend	small token of appreciation	Facebook specific
wall words signaling intimacy	aggregate of all below	not Facebook specific
» family words	see above	not Facebook specific
» swear words	see above	not Facebook specific
» friend words	see above	not Facebook specific
» sexual words	see above	not Facebook specific
» work words	see above	not Facebook specific
» leisure words	see above	not Facebook specific
» home words	see above	not Facebook specific
» money words	see above	not Facebook specific
» body words	see above	not Facebook specific
» religious words	see above	not Facebook specific
» health words	see above	not Facebook specific
» sum of all minus work words	see above	not Facebook specific
times tagged together in photo	same place/same time	not Facebook specific
appearances in photo	to contrast with above	not Facebook specific
posted photos of friend	to contrast with above	somewhat Facebook specific
friend posted photos of you	to contrast with above	somewhat Facebook specific
physical distance from friend	best category it fits into	not Facebook specific
married or in a relationship with friend		not Facebook specific
Parineral Somuras (Cennovattor)		
links posted to individual	giving information to other	not Escobook specific
application overlap		Facebook specific
links posted to all friends	broadcast to contrast	not Eacobook specific
initis posted to an menus	broadcast—to contrast	not racebook specific
Structural Holes Theory (Burt)		
overlap with friend's network	mutual friends	not Facebook specific
common group membership	topical interest groups	somewhat Facebook specific
common network membership	place & university networks	not Facebook specific
normalized overlapping interests	movies, music, etc.	not Facebook specific
normalized overlapping self-description	free text about oneself	not Facebook specific
number of non-overlapping communities	fan pages	not Facebook specific
various network metrics	see N in model formulation	not Facebook specific
common event participation	not possible	somewhat FB specific

Provision of Emotional Support (Wellman)

number of gifts

repeat

Facebook specific

wall positive emotional words wall negative emotional words inbox positive emotional words inbox negative emotional words

#### Social Class (Lin)

age difference education difference occupational difference difference in occupation ranks common group membership political difference religious difference LIWC + "b-day", emoticons from LIWC LIWC + "b-day", emoticons from LIWC not Facebook specific not Facebook specific not Facebook specific not Facebook specific

in reported ages in scale defined in Ch. 3 number of occupations use standard measure repeat in scale defined in Ch. 3 in scale defined in Ch. 3 not Facebook specific not Facebook specific not Facebook specific not Facebook specific somewhat Facebook specific not Facebook specific not Facebook specific

#### **General Demographics**

age	in years	not Facebook specific
sex	male of female	not Facebook specific
relationship status	one of limited choices	Facebook specific
looking for relationship?	yes or no	Facebook specific

#### **Usage Characterization**

number of applications installed
number of inbox messages
number of wall posts
number of notes
number of picture comments

apps which extend Facebook private communication activity measure activity measure activity measure Facebook specific somewhat Facebook specific not Facebook specific somewhat Facebook specific somewhat Facebook specific

# **APPENDIX B:**

# TIE STRENGTH SEMI-STRUCTURED INTERVIEW PROTOCOL

 Subject ID:
 \_\_\_\_\_\_

- 1. What factors influenced your decision to rate a friendship as strong?
- 2. Do you have strong friends that are not your Facebook friends?
- 3. If so, why not?
- 4. This is the ranking we have come up from your participation in the experiment. Do you feel that any of your friends are out of place? If so, why?
- 5. These are the variables that we have determined are most important for calling a friendship strong. Did you realize before the experiment that these would be most important?
- 6. How would you feel if a system (a computer, that is) analyzed your Facebook data to determine the strength of your friendships, and used it put prioritize information in your Newsfeed (the Facebook list of friend current events)?

# APPENDIX C: WE MEDDLE SEMI-STRUCTURED INTERVIEW PROTOCOL

Subject ID: \_\_\_\_\_ Date: \_\_\_\_\_

This protocol is semi-structured; interviewer may pose emerging questions (not listed below) as appropriate according to the interview conversation. Each interview is anticipated to last no more than 20 minutes.

# Part 1: Background

- 1. How long have you been a Twitter user?
- 2. Do you use it often?
- 3. Why would you say that you do it? In general, what are your motivations for using Twitter?

# Part 2: Specific to We Meddle

- 4. Do you recall using We Meddle? Did you make the lists? If so, did you modify them before using We Meddle? After outside of We Meddle?
- 5. Did the Inner Circle really reflect your inner circle?
- 6. Did the Outer Circle really reflect your outer circle?
- 7. Did any of We Meddle's guesses seem particularly out of place? Do you remember which ones? Can you tell me about those?
- 8. Did you consider removing anyone from the lists? Did you add anyone afterward?
- 9. Did you try the client?
- 10. If you care to do so, please give us any other feedback which you think might help develop this idea going forward.

# BIBLIOGRAPHY

Adamic, L. A. and Adar, E. (2003). Friends and neighbors on the Web. *Social Networks*, **25**(3), 211–230.

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

Albert, R. and Barabási, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, **74**(1), 47–97.

Archak, N., Ghose, A., and Ipeirotis, P. G. (2007). Show me the money!: deriving the pricing power of product features by mining consumer reviews. In *KDD '07*: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 56–65.

Backstrom, L., Huttenlocher, D., Kleinberg, J., and Lan, X. (2006). Group formation in large social networks: membership, growth, and evolution. In *KDD* '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, 44–54.

Barzilay, R. and Elhadad, M. (1997). Using lexical chains for text summarization. In Proceedings of the ACL Workshop on Intelligent Scalable Text Summarization, 10–17.

Bearman, P. S., Moody, J., and Stovel, K. (2004). Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks. *American Journal of Sociology*, **110**(1), 44–91.

Bederson, B. B. and Hollan, J. D. (1994). Pad++: a zooming graphical interface for exploring alternate interface physics. In *UIST '94*: Proceedings of the 7th annual ACM symposium on User interface software and technology, 17–26.

Benkler, Y. (2006). The Wealth of Networks: How Social Production Transforms Markets and Freedom. Yale University Press.

Berger, J. and Milkman, K. (2009). Social Transmission and Viral Culture. Preprint. Retrieved August 28, 2010 from http://papers.srn.com/sol3/papers.cfm?abstract\_id=1528077.

Bernard, H. R., Killworth, P., Kronenfeld, D., and Sailer, L. (1984). The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology*, **13**(1), 495–517.

Bernard, R. H., Killworth, P. D., and Sailer, L. (1979). Informant accuracy in social network data IV: a comparison of clique-level structure in behavioral and cognitive network data. *Social Networks*, **2**(3), 191–218.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. (2009). Network analysis in the social sciences. *Science*, **323**(5916), 892–895.

boyd, d. (2002). FACETED ID/ENTITY: Managing representation in a digital world. MIT Master's Thesis.

boyd, s. (2006). Friends, friendsters, and top 8: Writing community into being on social network sites. *First Monday*, **11**(12).

Burt, R. (2000). The Network Structure of Social Capital. *Research in Organizational Behavior*, **22**, 345–423.

Burt, R. S. (1987). Social Contagion and Innovation: Cohesion Versus Structural Equivalence. *The American Journal of Sociology*, **92**(6), 1287–1335.

Burt, R. S. (2004). Structural Holes and Good Ideas. *American Journal of Sociology*, **110**(2), 349–399.

Burt, R. (1995). Structural Holes: The Social Structure of Competition. Harvard University Press.

Butts, C. T. (2009). Revisiting the foundations of network analysis. Science, 325(5939), 414–416.

Castellano, C., Fortunato, S., and Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, **81**(2), 591+.

Centola, D. and Macy, M. (2007). Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, **113**(3), 702–34.

Chopra, S., Thampy, T., Leahy, J., Caplin, A., and Lecun, Y. (2007). Discovering the hidden structure of house prices with a non-parametric latent manifold model. In *KDD '07*: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 173–182.

Christakis, N. A. and Fowler, J. H. (2007). The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine*, **357**(4), 370–379.

Christakis, N. A. and Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, **358**(21), 2249.

Clark, H. H. (1993). Arenas of Language Use. University Of Chicago Press.

Cohen, S., Brissette, I., Skoner, D. P., and Doyle, W. J. (2000). Social Integration and Health: The Case of the Common Cold. *Journal of Social Structure*, **1**(3), 1–7.

Coleman, J., Katz, E., and Menzel, H. (1957). The diffusion of an innovation among physicians. *Sociometry*, **20**(4), 253–270.

Collins, C. (2006). DocuBurst: Document Content Visualization using Language Structure. In Proc. of *IEEE Symp. on Information Visualization (InfoVis 2006)*, Poster Session, 112–113.

comScore (2008). comScore Media Metrix Ranks Top 50 U.S. Web Properties for June 2008. Retrieved from http://www.comscore.com/press/release.asp?press=2339. Constant, D., Sproull, L., and Kiesler, S. (1996). The Kindness of Strangers: The Usefulness of Electronic Weak Ties for Technical Advice. *Organization Science*, 7(2), 119–135.

Cosley, D., Frankowski, D., Terveen, L., and Riedl, J. (2007). SuggestBot: using intelligent task routing to help people find work in wikipedia. In *IUI '07*: Proceedings of the 12th international conference on Intelligent user interfaces, 32–41.

Cunningham, W. H., Cunningham, I. C., and Green, R. T. (1977). The Ipsative Process to Reduce Response Set Bias. Public *Opinion Quarterly*, **41**(3), 379–384.

De Choudhury, M., Mason, W. A., Hofman, J. M., and Watts, D. J. (2010). Inferring Relevant Social Networks from Interpersonal Communication. In Proc. *WWW 2010:* Proceedings of the 19th international conference on World wide web, 301–310.

Dumais, S. T., Furnas, G. W., Landauer, T. K., Deerwester, S., and Harshman, R. (1988). Using latent semantic analysis to improve access to textual information. In *CHI* '88: Proceedings of the SIGCHI conference on Human factors in computing systems, 281–285.

Easley, D. and Kleinberg, J. (2010). Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press.

Esuli, A. and Sebastiani, F. (2006). SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining.

Farnham, S., Portnoy, W., Turski, A., Cheng, L., and Vronay, D. (2003). Personal map: Automatically modeling the user's online social network. Proceedings of *INTERACT'03*.

Festinger, L., Schachter, S., and Back, K. (1963). Social pressures in informal groups. Stanford University Press.

Fisher, D., Hogan, B., Brush, A. J., Smith, M., and Jacobs, A. (2006). Using Social Sorting to Enhance Email Management. In *Human-Computer Interaction Consortium (HCIC'06)*, Fraser, Colorodo.

Flake, G. W. and Lawrence, S. (2002). Efficient SVM Regression Training with SMO. Machine Learning, **46**(3), 271–290.

Fowler, J. H. and Christakis, N. A. (2008). Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study. *British Medical Journal*, **337**(dec04), a2338+.

Frakes, W. B. and Baeza-Yates, R. (1992). Information Retrieval: Data Structures and Algorithms. Prentice Hall.

Freeman, L. C. (1992). The Sociological Concept of "Group" An Empirical Test of Two Models. *The American Journal of Sociology*, **98**(1), 152–166.

Frenzen, J. K. and Davis, H. L. (1990). Purchasing Behavior in Embedded Markets. *The Journal of Consumer Research*, **17**(1), 1–12.

Friedkin, N. E. (1980). A Test of Structural Features of Granovetter's Strength of Weak Ties Theory. *Social Networks*, **2**(4), 411–422.

Friedland, L. and Jensen, D. (2007). Finding tribes: identifying close-knit individuals from employment patterns. In *KDD '07*: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 290–299.

Garg, N. and Weber, I. (2008). Personalized, interactive tag recommendation for flickr. In *RecSys* '08: Proceedings of the 2008 ACM conference on Recommender systems, 67–74.

Gergle, D., Kraut, R. E., and Fussell, S. R. (2006). The Impact of Delayed Visual Feedback on Collaborative Performance. In *Proc. CHI*, 1303–1312.

Gilbert, E. and Karahalios, K. (2009). Predicting tie strength with social media. In *CHI '09*: Proceedings of the 27th international conference on Human factors in computing systems, 211–220.

Gilbert, E., Karahalios, K., and Sandvig, C. (2008). The network in the garden: an empirical analysis of social media in rural life. In *CHI '08:* Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, 1603–1612.

Glaser, B. and Strauss, A. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Transaction.

Golder, S., Wilkinson, D. M., and Huberman, B. A. (2007). Rhythms of Social Interaction: Messaging Within a Massive Online Network. In *3rd International Conference on Communities and Technologies* (CT2007).

Granovetter, M. S. (1973). The Strength of Weak Ties. *The American Journal of Sociology*, **78**(6), 1360–1380.

Granovetter, M. (1974). Getting a Job: A Study of Contacts and Careers. University Of Chicago Press.

Granovetter, M. (1978). Threshold models of collective behavior. *The American Journal of Sociology*, **83**(6), 1420.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, 201–233.

Granovetter, M. (2005). The impact of social structure on economic outcomes. *Journal of Economic Perspectives*, **19**(1), 33–50.

Greif, I. (2009). Venturing into social software. In *CTS '09*: Proceedings of the 2009 International Symposium on Collaborative Technologies and Systems, xiv+.

Grudin, J. (1988). Why CSCW applications fail: problems in the design and evaluation of organizational interfaces. In *CSCW* '88: Proceedings of the 1988 ACM conference on Computer-supported cooperative work, 85–93.

Gruhl, D., Guha, R., Liben-Nowell, D., and Tomkins, A. (2004). Information diffusion through blogspace. In *Proc. WWW 2004*, 491–501.

Hampton, K. and Wellman, B. (2003). Neighboring in Netville: How the Internet supports community and social capital in a wired suburb. *City and Community*, **2**(4), 277–311.

Hansen, M. T. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits. *Administrative Science Quarterly*, **44**(1), 82–111.

Hargittai, E. (2007). Whose Space? Differences Among Users and Non-Users of Social Network Sites. *Journal of Computer-Mediated Communication*, **13**(1), 276–297.

Haythornthwaite, C. (2001). The strength and the impact of new media. System Sciences, 2001. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, 10+.

Haythornthwaite, C. (2002). Strong, Weak, and Latent Ties and the Impact of New Media. *Information Society*, **18**(5), 385–401.

Haythornthwaite, C. and Wellman, B. (1998). Work, friendship, and media use for information exchange in a networked organization. *J. Am. Soc. Inf. Sci.*, **49**(12), 1101–1114.

Heider, F. (1946). Attitudes and cognitive organization. Journal of Psychology, 21(1), 107-112.

Henderson Jr, D. A. and Myer, T. H. (1977). Issues in message technology. In *Proceedings of the fifth symposium on Data communications*.

Hidalgo, C. A., Klinger, B., Barabasi, A. L., and Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, **317**(5837), 482–487.

Hogan, B. J. (2009). Networking in Everyday Life. University of Toronto, PhD Thesis. Department of Sociology.

Hsieh, G. (2009). Don't ask me: designing social Q&A services. Crossroads, 16(2), 24-26.

Huberman, B. A., Romero, D. M., and Wu, F. (2008). Social networks that matter: Twitter under the microscope. *First Monday*, 14(1).

Iqbal, S. T. (2008). A Framework for Intelligent Notification Management in Multitasking Domains. University of Illinois, PhD thesis. Department of Computer Science.

Ivkovic, Z. and Weisbenner, S. (2007). Information Diffusion Effects in Individual Investors' Common Stock Purchases: Covet Thy Neighbors' Investment Choices. *Review of Financial Studies*, **20**(4), 1327–1357. Jackson, M. O. (2008). Social and Economic Networks. Princeton University Press.

Kahanda, I. and Neville, J. (2009). Using transactional information to predict link strength in online social networks. *Proc. ICWSM*, 2009.

Katz, E., Lazarsfeld, P. F., and Roper, E. (2006). Personal influence: The part played by people in the flow of mass communications. Transaction Publishers.

Kempe, D., Kleinberg, J., and Tardos, E. (2003). Maximizing the spread of influence through a social network. In *KDD '03*: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, 137–146.

Kittur, A. and Kraut, R. E. (2010). Beyond Wikipedia: coordination and conflict in online production groups. In *CSCW 2010*: Proceedings of the 2010 ACM conference on Computer supported cooperative work, 215–224.

Klovdahl, A. (2005). Social network research and human subjects protection: Towards more effective infectious disease control. *Social Networks*, **27**(2), 119–137.

Kossinets, G., Kleinberg, J., and Watts, D. (2008). The structure of information pathways in a social communication network. In *KDD '08*: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, 435–443.

Kossinets, G. and Watts, D. J. (2006). Empirical Analysis of an Evolving Social Network. *Science*, **311**(5757), 88–90.

Krackhardt, D. (1993). The Strength of Strong Ties: The Importance of Philos in Organizations. In N. Nohria and R. Eccles (Ed.), Networks and Organizations: Structure, Form, and Action (216–239). Boston, MA: Harvard Business School Press.

Krackhardt, D. and Stern, R. N. (1988). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly*, **51**(2), 123–140.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., and Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological wellbeing?. *American psychologist*, **53**, 1017–1031.

Kwak, H., Lee, C., Park, H., and Moon, S. (2010). What is Twitter, a social network or a news media?. In *WWW '10*: Proceedings of the 19th international conference on World wide web, 591–600.

Lampe, C., Ellison, N., and Steinfield, C. (2008). Changes in Participation and Perception of Facebook. In *CSCW 2008:* Proceedings of the 2008 ACM conference on Computer supported cooperative work, 721–730.

Lau, T. (2009). Venture Research. Communications of the ACM Blog, Retrieved from http://cacm.acm.org/blogs/blog-cacm/38421-venture-research/fulltext.

Laumann, E. O., Gagnon, J. H., Michaels, S., Michael, R. T., and Coleman, J. S. (1989). Monitoring the AIDS epidemic in the United States: a network approach. *Science*, **244**(4909), 1186–1189.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., and Van Alstyne, M. Computational Social Science. *Science*, **323**(5915), 721–723.

Leskovec, J., Adamic, L. A., and Huberman, B. A. (2007). The dynamics of viral marketing. *ACM Transactions on the Web*, **1**(1), 5.

Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010). Signed Networks in Social Media. In Proc. *CHI 2010:* Proceeding of the twenty-eigth annual SIGCHI conference on Human factors in computing systems, 1361–1370.

Levin, D. Z. and Cross, R. (2004). The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer. *Management Science*, **50**(11), 1477–1490.

Liben-Nowell, D. and Kleinberg, J. (2003). The Link Prediction Problem for Social Networks. In *CIKM '03:* Proceedings of the twelfth international conference on Information and knowledge management, 556–559.

Liben-Nowell, D. and Kleinberg, J. (2008). Tracing information flow on a global scale using Internet chain-letter data. *Proceedings of the National Academy of Sciences*, **105**(12), 4633–4638.

Lieberman, E. and Miller, R. C. (2007). Facemail: showing faces of recipients to prevent misdirected email. Proceedings of the 3rd symposium on Usable privacy and security, 122–131.

Lin, N., Dayton, P. W., and Greenwald, P. (1978). Analyzing the instrumental use of relations in the context of social structure. *Sociological Methods & Research*, 7(2), 149.

Lin, N., Ensel, W. M., and Vaughn, J. C. (1981). Social Resources and Strength of Ties: Structural Factors in Occupational Status Attainment. *American Sociological Review*, **46**(4), 393–405.

Macy, M. W. and Willer, R. (2002). From Factors to Actors: Computational Sociology and Agent-Based Modeling. *Annual Review of Sociology*, **28**(1), 143–166.

Marin, A. (2004). Are respondents more likely to list alters with certain characteristics?: Implications for name generator data. *Social Networks*, **26**(4), 289–307.

Marsden, P. V. (1987). Core Discussion Networks of Americans. *American Sociological Review*, **52**(1), 122–131.

Marsden, P. V. (1990). Network Data and Measurement. *Annual Review of Sociology*, **16**(1), 435–463.

Marsden, P. V. and Campbell, K. E. (1984). Measuring Tie Strength. Social Forces, 63(2), 482–501.

Mcpherson, M., Lovin, L. S., and Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, **27**, 415–444.

Meagher, C. (2008). Caught on the Web: Internet Used More and More Often in Criminal Investigations. The Santa Barbara Independent.

Milgram, S. (1967). The small world problem. Psychology Today, 2(1), 60-67.

Milgram, S. (1977). The familiar stranger: An aspect of urban anonymity. *The Individual in a Social World*, 51–53.

Neustaedter, C., Brush, A., Smith, M., and Fisher, D. (2005). The Social Network and Relationship Finder: Social Sorting for Email Triage. In Conference on E-mail and Anti-Spam.

Newman, M. E. J. (2006). Modularity and Community Structure in Networks. *Proceedings of the National Academy of Sciences*, **103**(23), 8577-8582.

Nowell, D. L. and Kleinberg, J. (2003). The link prediction problem for social networks. In *CIKM* '03: Proceedings of the twelfth international conference on Information and knowledge management, 556–559.

Olson, G. M. and Olson, J. S. (2000). Distance matters. *Human-computer Interaction*, **15**(2), 139–178.

Onnela, J. P., Saramaki, J., Hyvonen, J., Szabo, G., Lazer, D., Kaski, K., Kertesz, J., and Barabasi, A. L. (2006). Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, **104**(18), 7332–7336.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *EMNLP*, 79–86.

Parks, M. R. and Floyd, K. (1996). Making Friends in Cyberspace. *Journal of Computer-Mediated Communication*, **1**(4).

Pennebaker, J. W. and Francis, M. E. (1999). Linguistic Inquiry and Word Count. Lawrence Erlbaum.

Potterat, J. J., Muth, S. Q., Rothenberg, R. B., Zimmerman-Rogers, H., Green, D. L., Taylor, J. E., Bonney, M. S., and White, H. A. (2002, April). Sexual network structure as an indicator of epidemic phase. *Sexually Transmitted Infections*, **78**, i152–i158.

Proulx, S., Promislow, D., and Phillips, P. (2005). Network thinking in ecology and evolution. *Trends in Ecology & Evolution*, **20**(6), 345–353.

Rapleaf (2008). RapLeaf Study of Social Network Users vs. Age. Retrieved from http://business.rapleaf.com/company%5C\_press%5C\_2008%5C\_06%5C\_18.html.

Reagans, R. and Zuckerman, E. W. (2001). Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science*, **12**(4), 502–517.

Redner, S. (2008). Networks: Teasing out the missing links. Nature, 453(7191), 47-48.

Riloff, E. and Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In Proceedings of the 2003 conference on Empirical methods in natural language processing, 105–112.

Rogers, E. M. (1995). Diffusion of Innovations, Fourth Edition. Free Press.

Rothenberg, R. B., Potterat, J. J., Woodhouse, D. E., Muth, S. Q., Darrow, W. W., and Klovdahl, A. S. (1998). Social network dynamics and HIV transmission. *Aids*, **12**(12), 1529.

Ryan, B. and Gross, N. C. (1944). The diffusion of hybrid seed corn in two Iowa communities. *Ru-ral Sociology*, **8**(1), 15–24.

Salton, G., Singhal, A., Mitra, M., and Buckley, C. (1997). Automatic text structuring and summarization. *Information Process Management*, **33**(2), 193–207.

Schaefer, C., Coyne, J. C., and Lazarus, R. S. (1981). The Health-related Functions of Social Support. *Journal of Behavioral Medicine*, **4**(4), 381–406.

Scherer, C. W. and Cho, H. (2003). A Social Network Contagion Theory of Risk Perception. *Risk Analysis*, **23**(2), 261–267.

Shi, X., Adamic, L. A., and Strauss, M. J. (2007). Networks of Strong Ties. *Physica A: Statistical Mechanics and its Applications*, **378**(1), 33–47.

Siegel, D. A. (2009). Social networks and collective action. *American Journal of Political Science*, **53**(1), 122–138.

Smith, M. A. and Fiore, A. T. (2001). Visualization components for persistent conversations. In *CHI '01:* Proceedings of the SIGCHI conference on Human factors in computing systems, 136–143.

Strang, D. (1996). Inducing a network influence structure from multiple diffusion processes. Annual Meeting of American Sociological Association, New York.

Strang, D. and Soule, S. A. (1998). Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual Review of Sociology*, **24**(1), 265–290.

Sun, E., Rosenn, I., Marlow, C., and Lento, T. (2009). Gesundheit! Modeling Contagion Through Facebook News Feed. In *Proc. ICWSM*.

Tahmincioglu, E. (2008). Facebook Friends As Job References?. Retrieved September 14, 2008. http://www.msnbc.msn.com/id/26223330
Thomas, M., Pang, B., and Lee, L. (2006). Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In *EMNLP*, 327–335.

Uzzi, B. (1999). Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firms Seeking Financing. *American Sociological Review*, **64**(4), 481–505.

van Dongen, S. (2000). Graph Clustering by Flow Simulation. University of Utrecht. PhD Dissertation.

Viégas, F. B., Wattenberg, M., and Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. In CHI '04: Proceedings of the 2004 conference on Human factors in computing systems, 575–582.

Viégas, F. B., Wattenberg, M., van Ham, F., Kriss, J., and McKeon, M. (2007). ManyEyes: a Site for Visualization at Internet Scale. *IEEE Transactions on Visualization and Computer Graphics*, **13**(6), 1121–1128.

Viswanath, B., Mislove, A., Cha, M., and Gummadi, K. P. (2009). On the evolution of user interaction in facebook. Proceedings of the 2nd ACM workshop on Online social networks.

Wattenberg, M. (2002). Arc diagrams: visualizing structure in strings. In *Information Visualization*, IEEE Symposium on Information Visualization, 110–116.

Wattenberg, M. and Viegas, F. (2008). The Word Tree, an Interactive Visual Concordance. In *Information Visualization*, IEEE Symposium on Information Visualization.

Watts, D. J. (2004). Six Degrees: The Science of a Connected Age. W.W. Norton & Co.

Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., and Haythornthwaite, C. (1996). Computer networks as social networks: Collaborative work, telework, and virtual community. *Annual Review of Sociology*, **22**(1), 213–238.

Wellman, B. and Wortley, S. (1990). Different Strokes from Different Folks: Community Ties and Social Support. *The American Journal of Sociology*, **96**(3), 558–588.

Whittaker, S., Jones, Q., Nardi, B., Creech, M., Terveen, L., Isaacs, E., and Hainsworth, J. (2004). ContactMap: Organizing communication in a social desktop. *ACM Transactions on Computer-Human Interactaction*, **11**(4), 445–471.

Whittaker, S., Jones, Q., and Terveen, L. (2002). Contact management: identifying contacts to support long-term communication. In *CSCW '02:* Proceedings of the 2002 ACM conference on Computer supported cooperative work, 216–225.

Whittaker, S., Terveen, L., Hill, W., and Cherny, L. (1998). The Dynamics of Mass Interaction. In *CSCW '98:* Proceedings of the 1998 ACM conference on Computer supported cooperative work, 257–264.

Williams, D., Ducheneaut, N., Xiong, L., Zhang, Y., Yee, N., and Nickell, E. (2006). From tree house to barracks: The social life of guilds in World of Warcraft. *Games and Culture*, **1**(4), 338.

Wilson, T., Hoffmann, P., Somasundaran, S., Kessler, J., Wiebe, J., Choi, Y., Cardie, C., Riloff, E., and Patwardhan, S. (2005). OpinionFinder: a system for subjectivity analysis. In Proceedings of *HLT/EMNLP* on Interactive Demonstrations, 34–35.

Wuchty, S. (2009). What is a social tie?. *Proceedings of the National Academy of Sciences*, **106**(36), 15099–15100.

Xiang, R., Neville, J., and Rogati, M. (2010). Modeling Relationship Strength in Online Social Networks. In *WWW '10:* Proceedings of the 19th international conference on World wide web, 981– 990.

Zanardi, V. and Capra, L. (2008). Social ranking: uncovering relevant content using tag-based recommender systems. In *RecSys '08*: Proceedings of the 2008 ACM conference on Recommender systems, 51–58.