Computing and Building around Tie Strength in Social Media

Eric Gilbert
Georgia Institute of Technology,
Atlanta,
USA
gilbert@cc.gatech.edu

Karrie Karahalios
University of Illinois at Urbana-Champaign,
USA
kkarahal@illinois.edu
Editorial Scope

Topics

Foundations and Trends® in Human-Computer Interaction publishes surveys and tutorials on the foundations of human-computer interaction. The scope is broad. The list of topics below is meant to illustrate some of the coverage, and is not intended to be an exhaustive list.

- History of the research community
- Design and evaluation
- Theory
- Technology
- Computer supported cooperative work
- Interdisciplinary influence
- Advanced topics and trends

Information for Librarians

Foundations and Trends® in Human-Computer Interaction, 2013, Volume 7, 4 issues. ISSN paper version 1551-3955. ISSN online version 1551-3963. Also available as a combined paper and online subscription.
Computing and Building Around Tie Strength in Social Media

Eric Gilbert
Georgia Institute of Technology,
Atlanta,
USA
gilbert@cc.gatech.edu

Karrie Karahalios
University of Illinois at Urbana-Champaign,
USA
kkarahal@illinois.edu
# Contents

1 Introduction 3
   1.1 What is tie strength? 5
   1.2 Scops 6
   1.3 Existing approaches 7
   1.4 Scenarios 7
   1.5 Contributions 8

2 Literature Review 10
   2.1 Tie strength 11
   2.2 Socially-rendered social media 19

3 Computing Tie Strength 26
   3.1 Research questions 27
   3.2 What does one number mean? 28
   3.3 Method 29
   3.4 Participants 30
   3.5 Predictive variables 31
   3.6 Dependent Variables 36
   3.7 Statistical Methods 37
   3.8 Results 40
   3.9 When the model breaks down: High residuals 45

Full text available at: http://dx.doi.org/10.1561/1100000044
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.10 Asymmetric friendships</td>
<td>47</td>
</tr>
<tr>
<td>3.11 Confounding the medium</td>
<td>48</td>
</tr>
<tr>
<td>3.12 Discussion</td>
<td>49</td>
</tr>
<tr>
<td>3.13 Conclusion</td>
<td>55</td>
</tr>
<tr>
<td>4 Building Around Tie Strength</td>
<td>59</td>
</tr>
<tr>
<td>4.1 Research questions</td>
<td>60</td>
</tr>
<tr>
<td>4.2 The collapsed context problem</td>
<td>60</td>
</tr>
<tr>
<td>4.3 Social media or temporal media?</td>
<td>61</td>
</tr>
<tr>
<td>4.4 Twitter and tie strength</td>
<td>62</td>
</tr>
<tr>
<td>4.5 We Meddle</td>
<td>65</td>
</tr>
<tr>
<td>4.6 We Meddle lists</td>
<td>66</td>
</tr>
<tr>
<td>4.7 We Meddle’s client</td>
<td>70</td>
</tr>
<tr>
<td>4.8 Deployment</td>
<td>73</td>
</tr>
<tr>
<td>4.9 Does it generalize?</td>
<td>76</td>
</tr>
<tr>
<td>4.10 How users experienced we meddle</td>
<td>82</td>
</tr>
<tr>
<td>5 Next Steps &amp; Conclusions</td>
<td>92</td>
</tr>
<tr>
<td>5.1 Models and predictors of tie strength</td>
<td>92</td>
</tr>
<tr>
<td>5.2 Rendering social media socially</td>
<td>95</td>
</tr>
<tr>
<td>5.3 Conclusions</td>
<td>98</td>
</tr>
<tr>
<td>References</td>
<td>101</td>
</tr>
</tbody>
</table>
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.

Simply put, we do not understand a basic property of relationships expressed online. This monograph takes a wide view of the problem, merging the theories behind tie strength with the data from social media. We 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 article makes two core 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. We argue that it also offers us a chance to rethink many of social media’s most fundamental design elements. Next, we 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*, we 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. It may be fair to see this monograph 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. We conclude the article by reflecting on other ways design might appropriate ideas like tie strength in social computing.
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 monograph 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, we 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,
Introduction

sometimes they call someone hoping for reference, “only to find that you were just drinking buddies.”

Academics see it too. For decades, various social sciences have documented how different types of relationships affect individuals and organizations. In this line of research, relationships are measured in the currency of tie strength. Loose acquaintances, known as weak ties, can help a friend generate creative ideas or find a job. They also expedite the transfer of knowledge across workgroups. Trusted friends and family, called strong ties, can affect emotional health and often join together to lead organizations through times of crisis. 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 large-scale social phenomena follows suit. A link between actors either exists or not, with the relationship having few properties of its own. 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?”

In other words, tie strength is a blind spot. This monograph addresses this problem, merging the theories behind tie strength with the data from social media. We show how to reconstruct tie strength from digital
1.1. What is tie strength?

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 our participants cleverly suggested, consider remaking Facebook’s Newsfeed to get rid of “people from high school I don’t give a crap about.”

Sometimes, when we talk about social media, it is easy to think of it as a new internet fad. However, social media is old, as old as the internet perhaps. It is 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 [69].

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 1 billion 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?

1.1 What is tie strength?

This section is about tie strength. Tie strength is a diffuse concept: it refers to a sense of closeness with 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 your acquaintances? Mark Granovetter, who introduced the concept, had this to say about tie strength’s fuzzy definition: “Most of us can agree, on a rough intuitive basis, whether a given tie is strong, weak, or absent.”

While most studies initially define tie strength as a feeling towards another person, we usually see them operationalize it as a single, countable measure. Since these studies want to know about “real life,” we see interview questions like, “How many times have you talked in the last month?” and “How often do you chat about political and social issues?,” as the way tie strength is measured. For these studies, in these contexts, this makes some amount of sense. Asking participants how often they see each other, while probably misreported [11], seems like hard data against which we make claims — even if it is not what we really want to measure.

In this monograph, we return to the original intent of tie strength: how close we feel to the people in our lives. We have the hard data: in social media, as opposed to real life, every interaction is recorded. Here, we map that concrete, irrefutable data to the feeling of closeness. This approach to tie strength has big advantages. Most importantly, it probably generalizes. Imagine we encounter a new context we have not studied yet. We studied instant messaging in the workplace, but not how non-profits use email. We could start from scratch. Or, we could start from what we already know about interpersonal closeness.

1.2 Scope

Tie strength 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 [56]. Questions like “How do mixtures of strong and weak ties affect someone’s ability to succeed in the workplace?” commonly make appearances in these studies.

The focus of this monograph, however, is how tie strength expresses itself in social media on the internet. While we study the ways in which
1.3 Existing approaches
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 we 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?” At a high level, however, it may be fair to see this monograph as linking the online to the offline; that is, linking social media traces to how we experience relationships in the real world.

1.3 Existing approaches

We are not the first people to care about tie strength, or the first person to try to model it. Until now, however, we have used simple heuristics to estimate tie strength. We take a different approach here. 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 [40], possessing at least one mutual friend [129], recency of communication [98] and interaction frequency [51, 58]. 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 simple heuristic like “call it a strong tie if they message each other at least \( N \) times” good enough? If it is bad, how bad? Before this work, we did not know. However, from our data we 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 presented here, 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.)

1.4 Scenarios

If we could model tie strength, how could we actually use it? Could we build something around it? We demonstrate one potential application
in this monograph (a re-rendering of social streams via tie strength), but we can imagine many. 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 us imagine that the next time she logs in, she has 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 our profile, but not these others,” and so on. And, they are notoriously hard to manage. In fact, the blog TechCrunch quoted 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. We believe 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.

1.5 Contributions

This monograph makes the following two specific contributions:

1. A rich, high-accuracy and general way to reconstruct tie strength from digital traces. The model presented in here uses

\footnote{http://techcrunch.com/2010/08/26/facebook-friend-lists.}
1.5. Contributions

more than 70 carefully chosen, theoretically meaningful 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 we later explore with We Meddle. With We Meddle, we show that the model presented in Section 3 generalizes 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, this work also 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 [80]. Section 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. We 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 Section 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 we hoped makes social media a bit better. It is the first application we are aware of to put a calibrated relational model at the heart of its design. With it, we 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 5000 people from around the world used We Meddle with no coercion or payment; they used it because they thought they would find value in it. We report on its design, its architecture, the reaction it received on the web and on follow-up interviews with users.
References


References


References


References


References


Full text available at: http://dx.doi.org/10.1561/1100000044
References


References


References


References


