

How You Met Me

Lada A. Adamic
 University of Michigan
 Facebook
 ladamic@umich.edu

Thomas M. Lento
 Facebook
 tlento@fb.com

Andrew T. Fiore
 Facebook
 atfiore@fb.com

Abstract

A popular Facebook meme asks a user’s friends to recall how they met that user and then to paste the same query in their own status. We study the spread of this particular meme, which engaged millions of Facebook users, and the insights into relationship formation that the resulting compilation of answers provides. We describe the locations, relationships, and circumstances that contribute to formations of friendships that are represented on Facebook.

Introduction

Online social networks are conducive to the propagation of memes. Memes are self-replicating pieces of information that encourage anyone exposed to create more copies of them, and thus expose their social networks as well. Memes adapt to their environment. On Facebook, many propagate as links and images that can be shared. Some, however, consist only of text, with copy and paste instructions embedded in the text itself. Copy and paste memes cover many different topics, from raising awareness of human diseases and conditions, to disseminating warnings about real and imagined hacker threats, to relaying pieces of wisdom and humor. One meme stands out as generating unusually rich and useful data, by encouraging its human hosts to report how they met one another. This is the most popular variant of the meme:

Do any of us really know everybody on our friend list? Here is a task for you. I want all my fb friends to comment on this status about how you met me. After you comment, copy this to your status so I can do the same. You will be amazed at the results you get in 12 hours.

This type of data, pertaining to individual social ties, is typically laboriously obtained through surveys and interviews. In this paper we take advantage of the millions of responses this memory meme generated both to glean an understanding about how a meme focusing on social interaction propagates, and to study the origins of friendships which are now captured on Facebook.

Copyright © 2012, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Related Work

It has long been understood that social ties are shaped by the contexts under which they form, from locational proximity to organizations to family ties (McPherson, Smith-Lovin, and Cook 2001). This has been indirectly evident in recent studies of tie formation and prediction. Kossinets and Watts (2006) showed that a few simple foci, e.g., attending the same classes or sharing email contacts in common, could be predictive of new connections within the email network of a university. However, to our knowledge, there is no large-scale study pinpointing the origin of social ties.

Memes themselves have been a subject of study, especially their propagation and characteristics in online networked environments (Shifman and Blondheim 2010; Shifman and Thelwall 2009). Leskovec et al. (2009) generated a large-scale dataset covering many mainstream media outlets and blogs as they propagated the same or similar pieces of text. Because meme instances in this prior work were gathered over many different sources, only a small fraction of transmissions paths can be reliably inferred, i.e. if one source directly cites another source in relation to the meme. In contrast, due to the specific nature of the meeting meme, which encourages individuals to first comment on the original, before making their own copy, we are able to regenerate a large diffusion tree, precisely mapping the spread of the meme. One other study of email chain letters (Liben-Nowell and Kleinberg 2008) was able to trace long chains. However, the data was reconstructed from a few emails containing the chains, leaving much data missing, and the dynamics of the spreading process in question (Golub and Jackson 2010). Sun et al. (2009), traced the spread of users’ ‘fanning’ of Facebook pages through the Facebook newsfeed, but did not find evidence of large cascades. In contrast, in this paper we study a meme that spread predominantly as a single large cascade.

Data description

We collected the data by searching over a set of anonymized status updates that contained copy and paste instructions, e.g. “copy”, “paste”, “repost”, and also the strings “how you met me” or “how we met”. The first variant of the meme appears in July 2009, and begins with “Leave one memory of how you met me...” We start our analysis a year later,

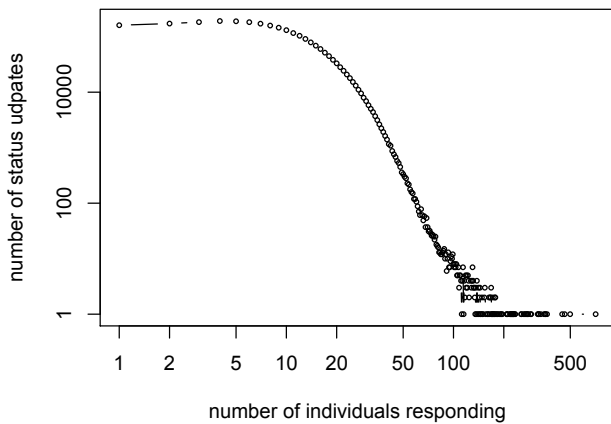


Figure 1: The distribution of the number of repliers per post.

in July 2010, concluding in November 2011. A manual examination of the resulting set of high occurrence status updates showed that they were nearly all variants of the memory meme. We detected 2,570,182 posts of the meme, with 24,199,921 comments, excluding self-comments typically posted by a user in response to someone recalling how they had met them. Figure 1 shows the distribution of the number of unique repliers for each post that was commented on. On average 7.5 distinct respondents commented on how they met the user.

User demographics

The meme appealed primarily to women, with 79.7% of posts being made by them. This is consistent with a sample of other copy-and-paste memes which averaged 73.3% female. The age distribution is broad (see Figure 2), with an average age of 34.5 ± 12.6 (excluding ages over 80, which are unlikely to be honestly reported). Since we are examining just the English variant of the meme, English-speaking countries are by far the most represented: (United States 70.1%, Great Britain 11.4%, Canada 5.8%, Australia 3.7%, New Zealand 0.7%, South Africa 0.6%, Ireland 0.5%).

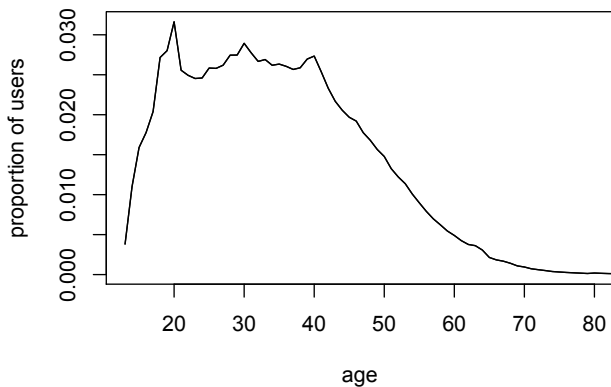


Figure 2: Age distribution of users propagating the meme.

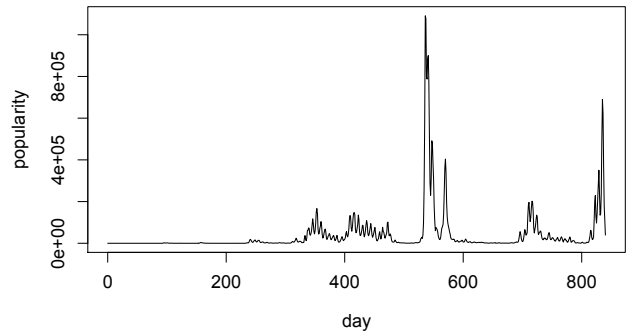


Figure 3: Popularity of the meme over time.

Meme variants

The main variants of the meme had four clauses that facilitated its spread. The first is the explicit copy and paste instruction without which the meme was unlikely to be viable. The second is the challenge and task nature of the activity. The meme raises doubt that the person reading it can remember how they met their friend, and then directly tasks the reader with the action. The third clause places an obligation of reciprocity on the reader: “*after you comment, copy this to your status so I can do the same*”. The fourth promises good results “*You will be amazed at the results you get in 12 hours*”.

Some variants apply further pressure: “*Now that you’ve read it, you must comment. DON’T CHEAT!!*” but were not as popular as the original. Some dropped the 12 hour clause, or promised great results without a fixed time frame. Others prefaced the meme with favorable reviews (“*This is fun!*” or “*This should be interesting...*”).

By far the most popular variants, excluding the original meme itself, twisted the meme around, asking friends to lie about how they had met:

I would like my Facebook friends to comment on this status, sharing how you met me. But I want you to LIE. That’s right, just make it up. After you comment, copy this to your status, so I can do the same. I bet half won’t read the instructions right!

Diffusion structure

The meme waxed and waned over a period of nearly 2 years, experiencing a few dramatic flare-ups, as shown in Figure 3. It is one of the most popular and long-lived Facebook status update memes during this period. Because the popularity of the meme showed such dramatic variation, and because it was nearly extinct for weeks at a time, it is not immediately obvious whether those engaged with the meme can be connected, or whether the meme spontaneously re-emerges as individuals decide to task their networks with recalling how they had met.

To trace the diffusion paths, we limited our analysis to those who had posted the meme to their status, and omitting those who had commented on the meme, but had not made the request of their own friends. We then time-ordered

the posts and comments, and identified the earliest post that a user had commented on as the parent of that user’s node in the diffusion tree. Tracing such ties in a breadth-first and directed manner, we found the longest chain to be 82 steps from root node to leaf node. We then recursively added all comment-to-post interactions that connected to that chain, and arrived at a giant component of 2,174,920 nodes, comprising over 84.4% of all the users we identified as having posted the meme. The graph is relatively sparse. Someone who had posted the meme had also commented on 2.3 posts on average.

Thus, even though the meme’s infection levels fell below 1000/day for weeks at a time, users participating in different major outbreaks that were spaced far apart in time were connected to one another through long infection chains. The giant component spanned the very beginning of our collection period (July 1, 2010) through the end (Nov. 19, 2011). It had reached its maximal depth of 82 on August 10, 2011. The disconnected components, the largest of which included 302 users, could have occurred for several reasons. A user may have copied the meme into their own status without commenting on someone else’s first. Or a post may have been deleted at a later point, leaving a gap in the chain. 4.9% of all nodes in the graph had no parent node.

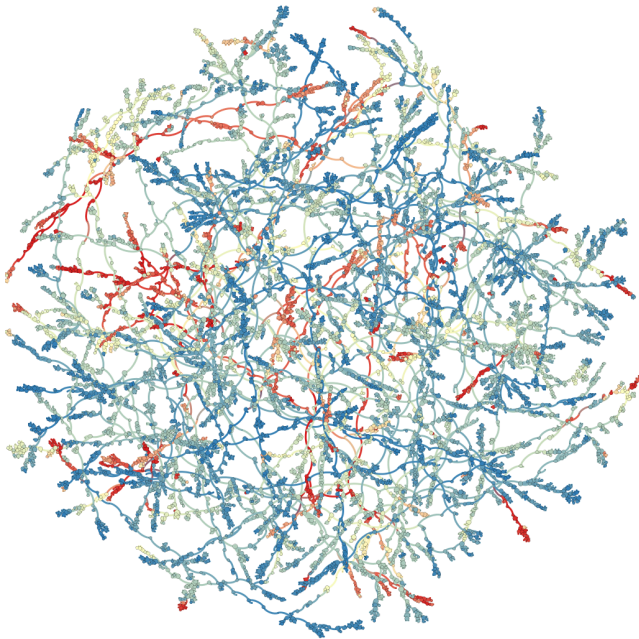


Figure 4: A brief portion of the diffusion of the meme. Nodes are users who posted their meme as their Facebook status. Edges are drawn between nodes if one user commented on the meme post of another. The colors denote the time at which the status update was posted, starting on the 6th of July 2010 (red), and ending 9th July 2010 (blue). The meme continued propagating past this point, eventually reaching millions of users.

Table 1: Friendship formation foci

category	keyword
school	school, class, grade, freshman...
work	coworker, work, job, training...
party	wedding, funeral, party...
birth	birth, born
church	church, choir, seminary...
neighbor	lived in the same, neighbor, next door...
sport	basketball, football, cheerleading
home	house, home, apartment...
hospital	hospital
student	my teacher, your teacher, my student

Origins of friendships

We analyzed the responses in the form of comments to the meme in order to gain an understanding regarding the origin of friendships. We used a series of simple regular expressions to extract different forms of responses: locations, intermediaries, direct familial relationships, introductions through others, and one-word answers. We omitted responses to posts that contained the word “lie”, as we did not want to include fabricated friendship starts. We also included just the first response of each commenter, as sometimes the asker replied to an initial response and spurred off-topic conversation. The data is naturally noisy, with some friends launching into lengthy recollections, while others were brief but did not stay to task. Some terms are also ambiguous, e.g. “orientation” can refer to an educational or a work setting. Furthermore, many responses contained just a single proper name, e.g. “Kennedy” to denote Kennedy High School, or Applebee’s, denoting either an employer or a hangout where the first encounter took place.

Despite the noisy nature of the data, the large volume of responses allows us to observe several patterns. A combination of regular expressions (excluding one-word responses) matched 42.9% percent of the responses. Another 2.9% were single word responses. Many of the unmatched responses were a mix of unique stories (e.g. “I was just chillin’ waiting on a sibling to show up, and I got you!!!”, jokes (e.g. “Do I know you lol?”), and non-response banter (e.g. “It’s been so long ago I don’t remember!”).

Foci

A few simple terms, partially listed in Table 1, were used to extract activity and geographic foci from 39.7% responses. The relative proportions of responses are a product of both the actual likelihood of meeting a friend in a specific environment and the keywords we used, although we tried to include as many of the frequently occurring words as possible. For example, many responses listed employers rather than mentioning the word “work” in their response. We observe that school is a popular milieu of friendship formation, but that it drops off slightly in favor of work as a function of age, suggesting an attrition in friends from school over time. Most mentions of hospitals appeared to pertain birth events, e.g. a mother telling her child that she met her at the hospital.

Although statistically significant ($\chi^2 > 10^4, p < 10^{-10}$), gender effects in friendship foci were mild, with two exceptions. Men were 57% more likely to meet a friend through sports than women, while women were 34% more likely than men have befriended a neighbor.

Connectors

Responses listed individuals through whom the connection had been made. We matched words such as *thru, through, dating, dated, married* followed by words such as *my, your, our mutual* as well as possessive words like “my” at the beginning of a comment. We recorded the next word that occurred, excluding adjectives. 4.2% of all responses were a match. Many listed family members as being the ones who made the introductions. The most frequent connectors were siblings, with an expected symmetry in the replies. “My brother” (1.6%) and “your brother” (1.4%) were the two most common responses, followed by “my sister” and “your sister”, for both men and women. Due to the freeform nature, there were many less frequent variants e.g. “my sis” (0.2%) and “ur brother” (0.2%). Following siblings, there were moms, dads, sons and daughters, as well as cousins.

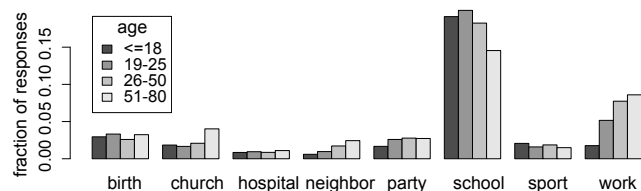


Figure 5: Locations and foci where friendships are formed, separated by age group. Age is a significant factor in where friendships originated ($\chi^2 > 10^5, p < 10^{-10}$).

Direct connections

Direct connections included phrases such as “I’m your...”, “You’re my”, or “We’re” and accounted for 0.3%, 1.4% and 0.3% of the responses, respectively. Direct connections, at least for family members, are often not expressed, because the commenter instead mentions the location, e.g. “in the hospital” or time “when you were born”. For the many direct relationships that were reported, their frequencies are as follows. Men were reported as being a cousin, brother, uncle, and nephew, followed cuz, dad, bro and son. For women, the responses were mirrored. Cousin was again the most frequent, followed by sister, aunt, niece, cuz, mom, auntie, and neighbor.

One word responses

One word responses accounted for 2.9% of the total and were a mix of the above categories, with “family”, and “cousin” being the most common. Other frequent one-word responses included Facebook – implying that the two individuals met online through the social networking site – and employers such as Walmart.

Conclusion

In this short paper we presented the spread of and content induced by a remarkable single meme. The meme reached millions of individuals and propagated in long chains. The responses it evoked from these individuals tell us that a significant fraction of Facebook friendships originated at school, even for individuals who have not been in school for decades. Interestingly, we fail to find a gender effect in introductions made through family: brothers are as likely to have made introductions for women as their sisters. The data is far richer than what this initial analysis has yielded, and we plan to explore it further in future work.

The richness of the data calls for a more in depth study of the properties of friendship formation. For example, the results suggest that different types of shared attributes, such as affiliations with schools and employers, can differ in their ability to predict friendships for different periods in an individual’s life and may vary by gender. There may also be some close ties, e.g. siblings and spouses, who play a disproportionate role in shaping individuals’ social networks in ways that have not been studied on a large scale. There is also the broader question of whether certain nodes in the network literally play the role of connectors by initially introducing many friends to one another. All these questions are interesting avenues for future work.

Acknowledgments

This work was supported in part by NSF IIS-0746646.

References

- Golub, B., and Jackson, M. 2010. Using selection bias to explain the observed structure of internet diffusions. *Proceedings of the National Academy of Sciences* 107(24):10833.
- Kossinets, G., and Watts, D. 2006. Empirical analysis of an evolving social network. *Science* 311(5757):88.
- Leskovec, J.; Backstrom, L.; and Kleinberg, J. 2009. Meme-tracking and the dynamics of the news cycle. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 497–506. ACM.
- 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.
- McPherson, M.; Smith-Lovin, L.; and Cook, J. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* 415–444.
- Shifman, L., and Blondheim, M. 2010. The medium is the joke: Online humor about and by networked computers. *New Media & Society* 12(8):1348.
- Shifman, L., and Thelwall, M. 2009. Assessing global diffusion with web memetics: The spread and evolution of a popular joke. *Journal of the American Society for Information Science and Technology* 60(12):2567–2576.
- Sun, E.; Rosenn, I.; Marlow, C.; and Lento, T. 2009. Gesundheit! modeling contagion through facebook news feed. *Proceedings of the Third International Conference on Weblogs and Social Media*.