Indifferent Attachment: The Role of Degree in Ranking Friends

David Liben-Nowell, Carissa Knipe, and Calder Coalson
Department of Computer Science, Carleton College, Northfield, MN 55057
{dlibenno, knipec, coalsonc}@carleton.edu

***Abstract*—Each user of the MySpace social network can designate a small subset of her friends as *Top Friends*, placing them in a rank-ordered list displayed prominently on her profile. By examining users' #1 (best) and #2 (second-best) friends, we discover that MySpace users are nearly indifferent to these two friends' popularities when choosing which to designate as their best friend. Other pairs of ranks (e.g., #1-vs.-#3, #2-vs.-#3, ...) also reveal no marked preference for a popular friend over a less popular one. To the extent that ranking decisions form a window into broader decisions about whom to befriend at all, these observations suggest that positing individuals' tendency to attach to popular people—as in network-growth models like preferential attachment—may not suffice to explain the heavy-tailed degree distributions seen in real networks.**

Introduction: choosing among friends. Different social relationships have different priorities. Implicitly or explicitly, we choose one friend over another: we answer *that* email first; we find time for coffee with *this* person after telling another we were too busy. The way we prioritize one friend over another is an interesting, and important, question about our social relationships; it reflects the way in which a social network is used and constructed. And our prioritization decisions may also offer a window into friend-making in general.

Computational research on social interactions has blossomed recently, though the preponderance of this work treats ties as binary: we are friends; we are not friends. But the importance of tie strength, and indeed the importance of *weak* ties, has long been studied by social scientists—most prominently in Mark Granovetter's "The Strength of Weak Ties" [7]. An expanding body of computational research explores relationship strength, too; a nonexhaustive sampling includes work on mobile phone users [11], trust in CouchSurfing [12], self-reported vs. behavioral tie strength in email [13]; "attention" across Facebook friends [1]; and predictive models of perceived tie strength built using profile similarity and measures of interaction [6, 14]. Perhaps most similar to our work is Kahanda and Neville's effort to classify edges as "strong" or "weak" (as represented by a friend's presence/absence in a Facebook app listing close friends) using the same type of structural and transactional properties [8].

Here, we address a fine-grained question of prioritization: rather than the *rating* of friendships on an absolute scale ("distant" to "close"), we consider *ranking* of friendships on a relative scale, using MySpace's *Top Friends.* Each MySpace user may designate k of her friends as "Top Friends," ranking them #1 through #k. (The user chooses the cardinality k; nearly all users choose $k \leq 40$, and $k = 8$ is common.) Friendship ranking—unlike rating or friendship existence—is a setting of scarce resources; after all, Alice can only have one #1-ranked friend. Accepting a distant acquaintance's friend request bears minimal cost—a mild reputational risk and mental strain for tracking the relationship—as does "grade inflation" in rating one's friends (seen in CouchSurfing [12], e.g.). But the scarcity of high friendship slots means that the ranking environment may shed a different light on potentially awkward social decisions. Some users work hard to avoid publicly declaring the ranking of their friendships [3], but enough MySpace users do provide rankings of real profiles that we can begin to pose, and answer, some interesting structural questions.

Thanks to Peter DeScioli, Elizabeth Koch, Robert Kurzban, Dave Musicant, Jeff Ondich, and Aaron Swoboda. This work was supported in part by NSF grant CCF-0728779 and by Carleton College.

Preference for popularity? We are most interested in a user's choice as to which of two individuals will be her #1-ranked friend and which will be her #2-ranked friend [5]. (We also consider #i-versus-#j decisions for all other $i, j \in \{1, \ldots, 8\}$.) Previous work [5] has shown that a user u has preference for *homophily* [9], and for those friends who rank u better [4]. But when Alice chooses how to prioritize two friends, does she prefer the popular Bob or the unpopular Charlie? Intuitively, Charlie is a more "committed" friend because he has fewer distractions, but Bob is a more "valuable" friend because he knows more people. Which of these competing intuitions affects Alice's decision more?

Consider a MySpace user u who names a best friend v_1 and a second-best friend v_2. Erase the ranking labels from the edges from u to v_1 and v_2; our central task is to predict which of $\{v_1, v_2\}$ is u's #1 friend. We study predictors that are based on the popularity (i.e., degree) of $\{v_1, v_2\}$, using an $\approx$11M-user sample of MySpace from 2007–2008 [5]. Here, we concentrate on three measures of popularity for a user v: (1) *total indegree*, the number of the $\approx$11M MySpace users who list v as a friend; (2) *rank$_k$ indegree*, the number of in-sample users who list v as their kth-best friend, and (3) *ranked indegree*, given by $\sum_k$(rank$_k$ indegree of v). These quantities all exhibit heavy-tailed distributions; any two of these measures (including rank$_{1\ldots8}$) are strongly positively correlated (and, with one minor exception, rank$_j$ and rank$_k$ indegree are more strongly correlated as $|j - k|$ gets smaller).

We filter the $\approx$11M profiles to identify $\approx$1.36M focal individuals u for which u's #1- and #2-ranked friends also appear in the sample. For each measure μ of popularity, we count the users u whose #1-ranked friend is more popular under μ than u's #2-ranked friend (a *win* for μ); the users whose #1- and #2-ranked friends are equally popular (a *tie*); and the users whose #2-ranked friend is more popular (a *loss*).

(We always compute popularity ignoring the edge from u; including this edge would "bake in" the right answer.)

Every popularity-based predictor has a success rate ($\frac{\text{wins}}{\text{wins+losses}}$) within 0.014 of 0.5—i.e., random guessing. Measured by deviation from 50%, the two most successful are rank_1 (51.4%) and rank_2 (49.3%) indegrees: there is a mild tendency for the #1-ranked friend to be ranked #1 more often by others, and for the #2-ranked friend to be ranked #2 more often by others. All other predictors perform between 49.5% and 50.2%. (Each μ has $n \geq 732414$ data points; viewing each μ as an n-trial binomial distribution, the standard error for each μ is ≤ 0.0012.) Even the most informative measures give only weak information, and indeed the various measures of popularity even differ in directionality: 7 of the 10 predictors say that individuals (weakly) prefer others who are *less* popular. For comparison, the geographic distance predictor ("u prefers the friend who is geographically closer to u") has success rate 0.564 and the relative rank predictor ("u prefers the friend who ranks u better") has success rate 0.689 [5].

The qualitative pattern of the #1-versus-#2 comparison remains true for other pairs of ranks i and $j > i$. The fraction of individuals who in the #i-versus-#j decision prefer the more popular candidate friend is generally close to or below 0.50, and only rarely greater than half. Generally, the fraction of individuals whose best friend is more popular than their #j-friend decreases with j. We do see the hints of one exception to this trend: out of two candidate friends of an individual u, the one with a higher rank_i indegree *is* generally (slightly) more likely to be u's $\#i$ friend. (That is, some users are "particularly good #k friends" for a single fixed value of k.)

Discussion: indifferent attachment? A number of network-growth models—most prominently, *preferential attachment (PA)* [2]—are designed to account for the heavy-tailed degree distribution of real social networks based on a "preference for popularity": that is, given a choice between two candidate friends, the more popular candidate tends to be preferred. While the basic form of PA (see [2] for details) does not speak directly to rankings of friends, the underlying preference for popularity does. PA can be most straightforwardly adapted to the ranked setting by modeling a node as ranking its neighbors in the order in which edges form. (So the #1-ranked friend for u is the friend u chose when joining the network; u's #2-ranked friend is the first node $v \neq u$ that chose u when v joined the network; etc.) We simulated this Ranked-PA (RPA) network growth model, and observed that friend ranking is much better predicted by popularity in RPA than in MySpace; over 95% of RPA nodes had a best friend who was more popular than their second-best friend, and between 59% and 61% of nodes had a #i-ranked friend more popular than their #$(i+1)$-ranked friend for $i \in \{2,3,4,5\}$. (In PA/RPA, the age of a node and the node's degree are positively correlated; thus the edge formed earlier is more likely to have been formed from a neighbor that would eventually become more popular. The #1-ranked friend is special, because it was chosen with explicit preference towards high degree instead of by age.)

The empirical and modeling observations of Barabási and Albert [2] sparked a large body of literature, empirical and theoretical—particularly as a hypothesis of the origin of the apparently ubiquitous heavy-tailed degree distributions. But our MySpace results, coupled with the simulations of RPA, suggest that a preference for popularity may not provide a full explanation for empirically observed heavy-tailed degree distributions: when a user is choosing which of two friends she prefers, the popularity of the two candidates is at best essentially uninformative about which will be chosen by that user, and at worst she actually prefers the less popular friend.

Perhaps the analogy between friendship existence choices and ranking choices is a weak one; perhaps people choose "to friend or not to friend" fundamentally differently from the way they choose best friends (and thus our results do not speak to unranked friendship decisions). It is an interesting direction for future research to assess this analogy's strength; an understanding of why, and how, a pair of individuals decide to assign a "friend" label to their relationship is generally missing from the computational literature on social networks—and this understanding is obviously crucial to properly interpreting the edges. But to the extent that the analogy holds, our observations suggest that network-growth models based on a preference for popularity miss important behavioral properties; we will need a different explanation to account for empirically observed heavy-tailed degree distributions. And even if ranking and befriending decisions are fundamentally different, the heavy-tailed degree distribution for rank_1 indegree seems to require an explanation fundamentally different from PA. Mitzenmacher [10] argues compellingly that, as a research community, we must move toward the necessary future direction of *validation* (or, at least, invalidation) in research on heavy-tailed degree distributions. We hope that the present work can serve as a small step forward in that enterprise.

References

[1] L. Backstrom, E. Bakshy, J. Kleinberg, T. Lento, and I. Rosen. Center of attention: How Facebook users allocate attention across friends. In *Proc. ICWSM*, July 2011.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[3] D. Boyd. Friends, Friendsters, and MySpace Top 8. *First Monday*, 11(12), December 2006.

[4] P. DeScioli and R. Kurzban. The alliance hypothesis for human friendship. *PLoS One*, 4(6):e5802, 2009.

[5] P. DeScioli, R. Kurzban, E. Koch, and D. Liben-Nowell. Best friends: Alliances, friend ranking, and the MySpace social network. *Perspectives on Psychological Science*, 6(1):6–8, January 2011.

[6] E. Gilbert and K. Karahalios. Predicting tie strength with social media. In *Proc. CHI*, April 2009.

[7] M. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, 1973.

[8] I. Kahanda and J. Neville. Using transactional information to predict link strength in online social networks. In *Proc. ICWSM*, May 2009.

[9] M. McPherson, L. Smith-Lovin, and J. Cook. Birds of a feather: Homophily in social networks. *Annual Rev. Sociol.*, 27:415–444, 2001.

[10] M. Mitzenmacher. Editorial: The future of power law research. *Internet Mathematics*, 2(4):525–534, 2006.

[11] J. Onnela, J. Saramaki, J. Hyvonen, G. Szabo, D. Lazer, K. Kaski, J. Kertész, and A.-L. Barabási. Structure and tie strengths in mobile communication networks. *PNAS*, 104(18):7332–7336, May 2007.

[12] C-Y. Teng, D. Lauterbach, and L. Adamic. I rate you. You rate me. Should we do so publicly? In *Proc. WOSN*, June 2010.

[13] S. Wuchty and B. Uzzi. Human communication dynamics in digital footsteps: A study of the agreement between self-reported ties and email networks. *PLoS One*, 6(11):e26972, 2011.

[14] R. Xiang, J. Neville, and M. Rogati. Modeling relationship strength in online social networks. In *Proc. WWW*, April 2010.