Loosing "Friends" on Facebook

Daniele Quercia
The Computer Laboratory
University of Cambridge
United Kingdom
dq209@cl.cam.ac.uk

Mansoureh Bodaghi
The Computer Laboratory
University of Cambridge
United Kingdom
mb794@cam.ac.uk

Jon Crowcroft
The Computer Laboratory
University of Cambridge
United Kingdom
jac22@cam.ac.uk

ABSTRACT

Recent analyses of self-reported data (mainly survey data) seem to suggest that social rules for ending relationships are transformed on Facebook. There seem to be a radical difference between offline and online worlds: reasons for ending online relationships are different than those for ending offline ones. These preliminary findings are, however, not supported by any quantitative evidence, and that is why we put them to test. We consider a variety of factors (e.g., age, gender, personality traits) that studies in sociology have found to be associated with friendship dissolution in the real world and study whether these factors are still important in the context of Facebook. Upon analyzing 34,012 Facebook relationships, we found that, on average, a relationship is more likely to break if it is not embedded in the same social circle, if it is between two people whose ages differ, and if one of the two is neurotic or introvert. Interestingly, we also found that a relationship with a common female friend is more robust than that with a common male friend. These findings are in line with previous analyses of another popular social-networking platform, that of Twitter. All this goes to suggest that there is not much difference between offline and online worlds and, given this predictability, one could easily build tools for monitoring online relations.

Author Keywords
social networks, tie decay, facebook, personality

ACM Classification Keywords
H.5.0. Information Interfaces and Presentation: General

General Terms
Human Factors; Measurement.

INTRODUCTION

In 2010, the *Oxford English Dictionary* named "unfriend" the word of the year, defining it as the act of removing "(someone) from a list of friends or contacts on a social networking site" [16].

The act of unfriending has been recently studied in the context of Twitter in quantitative ways. Kivran-Swaine *et al.* explored how breaking Twitter ties is related to sociological concepts such as strength of ties, embeddedness, and status [10]. At the same time, Kwak *et al.* separately found that the major factors that affect the decision to break a social relation (to 'unfollow', in Twitter parlance) are: relationship reciprocity, relationship duration, one's informativeness, and shared relationships [11].

In the context of Facebook, instead, studies on unfriending have been mostly qualitative or based on self-reported data. Sibona and Walczak have administer surveys to 690 Facebook users and found six main reasons for which one unfriends someone else on Facebook [20]. Four reasons are linked to online interactions (unimportant/frequent posts, polarizing posts, inappropriate posts and everyday life posts), and two are linked to offline interactions (disliked behavior and changes in the relationship). On a larger scale, few months ago, NM Incite, a Nielsen McKinsey company, has surveyed 1,865 adults in the United States and found a number of reasons people remove others from their friend lists, the most important of which (for 55% of the participants) was posting offensive comments, the second most imporant was not knowing someone very well, and the third was "they were trying to sell me something" [6]. These findings support views similar to that hold by cultural critic Lee Siegel. He has argued that the social rules for ending relationships are transformed online: "the act of unfriending acknowledges that the definition of friend is different from the traditional" [17].

Self-reported data seems to suggest that Facebook relationships end for reasons different than those associated with dissolution of real-life relationships. Such an assertion has, however, never been quantitatively tested. We thus set out to test whether factors associated with dissolution of real-life relationships still hold on Facebook and, in so doing, we make the following contributions:

- Based on literature review, we determine the most important predictors of real-life tie decay and consequently inform the design of our research (Section "Persistent and Broken Ties"). We find that these predictors are social embeddedness (i.e., whether two individuals' social circles are very similar), age difference, gender of common friends, and personality traits.

- We consider 34,012 Facebook relationships and study whether their decays are impacted by the four factors (Sections "Method" and "Results"). We find that a relationship is likely to break if it is not embedded, if it is between two users with a considerable age difference, and one of the two individuals is neurotic or introvert. We also find that

a relationship between two individuals having a common female friend is more robust than that between two individuals having a common male friend.

The main motivation for this kind of research is that, since both professional and private networks are moving online and are becoming virtualized, it is important to understand how individuals' conceptualization of social ties is being transformed by the Internet. It turns out that this alleged transformation is minimal and the fundamental social rules for ending relations are unaffected, and that points to the ease with which one could build social-networking tools for monitoring professional and private relationships (Section "Discussion").

PERSISTENT AND BROKEN TIES

Previous studies have found that the tendency of a relationship to decay and disappear is associated with the following four main factors:

Embeddedness. If two individuals have common friends, then their relationship is said to be socially embedded [5]. The more embedded a relationship, the more persistent [12]. That is largely because embeddedness and homophily go hand in hand [18]: embedded relationships tend to preferentially exist between similar individuals.

Gender Composition of Triads. In 1981, Bell conducted seminal work on the relationship between friendship and gender and went so far to claim that: "there is no social factor more important than that of sex in leading to friendship variation" [1]. He then commented: "Men turn more to women for close friendship" and described this as an emerging "new pattern" in cross-gender friendships. After Bell's work, research has consistently shown that women do indeed tend to describe their friendships in terms of emotional attachment and closeness and tend to make a deep commitment to their friends [19]. More recently, Kirk analyzed gender clustering in a network of teenagers using exponential random graph modelling (*ERGM*) and found that relationships between individuals having common female friends tend to be more persistent than those having common male friends [9].

Similarity. Homophily (i.e., "love of the same") "is the tendency of individuals to associate and bond with similar others"[1]. This tendency causes people's personal networks to be homogeneous with regard to a variety of sociodemographic, behavioural, and interpersonal characteristics. Homophily in ethnicity creates the strongest divides in our personal environments, followed by age, religion, education, occupation, and gender [13]: homophilous relationships are generally more sustainable, not least because people who are similar to each other are more likely to have repeated opportunities of interaction.

Personality. Personality traits predict a number of real-world behaviors. The five-factor model of personality, or the big five, is the most comprehensive, reliable and useful set of personality concepts [2, 4]. An individual is associated

with five scores that correspond to the five main personality traits and that form the acronym of *OCEAN*: imaginative, spontaneous, and adventurous individuals are high in **O**peness; ambitious, resourceful and persistent individuals are high in **C**onscientiousness; individuals who are sociable and tend to seek excitement are high in **E**xtraversion; those high in **A**greeableness are trusting, altruistic, tender-minded, and are motivated to maintain positive relationships with others; and emotionally liable and impulsive individuals are high in **N**euroticism. Individuals with stable relationships are generally those high in Extraversion and low in Neuroticism [7, 8].

To recap, this brief literature review suggests that the probability of a relationship to break decreases if the two individuals: 1) have common friends; 2) have at least one common female friend; 3) are similar for certain socio-demographic characteristics (e.g., age, occupation, education); and 4) are extrovert or emotionally stable.

METHOD

To test the extent to which those predictors hold in the context of Facebook, we gather data on persistent and broken relationships using a Facebook application.

Dataset

The Facebook application is called myPersonaltity, and users install it to take a variety of genuine personality and ability tests.

The resulting quality of the responses (quality of test results) is high: the scales' reliabilities are on average higher than reported in test, and the discriminant validity (average $r = .16$) is better than the ones obtained using traditional samples (average $r = .20$). That is largely because: 1) protocols that may be a product of inattentive, language incompetent, or randomly responding individuals are removed; and 2) appropriate incentives are in place - users are not paid to "be part of the study" and are solely motivated by the prospect of receiving reliable personality test results. Also, myPersonality users can give their consent to share their personality scores and profile information, and around 40% of them choose to do so. For those users, we collect their contact lists and end up with a dataset containing tuples in the form $<userID,time,\{friends\}>$, where $userID$ is the user's identifier in Facebook, $time$ is the time at which the user's list of contacts was crawled, and $\{friends\}$ is the list of the user's contacts. The dataset spans over a period of two and a half years from the beginning of 2009 to the first half of 2011.

Our goal is to identify, during the period of study, which Facebook social relationships persist and which relationships break. To this end, we split the dataset into one-month snapshots and, for the resulting 28 snapshots, we classify relationships as persistent or broken solely based on the only user activity data we have (i.e., the tuples $<userID,time,\{friends\}>$): a relationship is broken if it disappears after a certain snapshot, while it is persistent if it lasts from its appearance up to the final temporal snapshot. We consider an equal number of persistent and broken relationships and end up with a total of 34,012 relationships among
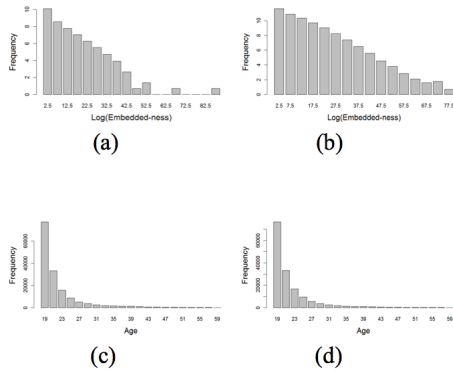
---
[1]http://en.wikipedia.org/wiki/Homophily

Figure 1. Frequency distributions of: (a) tie embeddedness for broken relationships; (b) tie embededdeness for persistent relationships; (c) age of individuals with broken ties; (d) age of individuals with persistent ties.

Figure 2. The relationship between probability p_{break} of breaking a relationship and: (a) the relationship's $embeddedness_{ij}$; (b) age difference between the two individuals in the relationship.

60% of women and 40% of men with a median age of 23. In Section "Discussion", we will explain why this sample is representative of the wider Facebook population in terms of age, gender, and number of social contacts.

RESULTS

Having this dataset at hand, we are now ready to test to which extent the act of unfriending is associated with the four factors we have identified in our literature review.

Embeddedness. We measure the embeddedness of a tie using Onnela *et al.*'s original formulation [15] :

$$embeddedness_{ij} = \frac{n_{ij}}{((k_i - 1) + (k_j - 1) - n_{ij})},$$

where n_{ij} is the number of common friends the two nodes (users) v_i and v_j have, and k_i (k_j) is the degree of node v_i (v_j). If v_i and v_j have no common friends, then we have $embeddedness_{ij} = 0$. We apply this formula to our equal number of broken and persistent relationships, and Figures 1(a) and 1(b) show the frequency distribution of the logarithm of embeddedness for both persistent and broken relationships. We take the logarithm of embeddedness because its distribution, as one can see, is skewed. We then plot probability p_{break} of a tie breaking as a function of the logarithm of its embeddedness. In Figure 2(a), we see that the more socially embedded a tie, the less likely it will break. We finally run a linear regression between p_{break} and a tie's embeddedness. The regression coefficient is statistically significant ($p < 0.001$) and negative, and the corresponding R^2 is as high as 0.72. These results suggest that the more embedded a tie, the lower its probability of breaking.

Gender Composition of Triads. Kirk showed that, in the offline context, triadic closure is stronger for triads in which there was at least one female [9]. To test this hypothesized correspondence, among our broken and persistent ties, we consider those ties that are embedded in triadic relationships (i.e., ties with common friends) and count the number of times the common friend is female or male. For triads with no common female friend, there are more broken relationships (54.5%) than persistent ones. By contrast, for triads with 2 females, the fraction of broken relationships goes down to
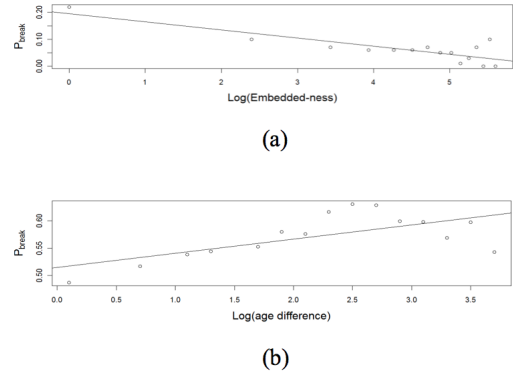
33.8%. A *t-test* on the means of the two distributions of female for persistent relationships and for broken relationships further confirms that persistent relationships are more likely to have a common female friend ($p < 0.05$).

Age difference. We plot probability p_{break} of a tie breaking as a function of the absolute age difference between the two individuals. In Figure 2(b), we see that the lower the age difference between two individuals, the less likely their tie will break. A linear regression between the probability of breaking a tie and the age difference in it returns a positive statistically-significant coefficient ($p < 0.01$), and the corresponding R^2 is high - it is 0.44.

Personality Traits of Extraversion and Neuroticism. For our users, we have their personality test results and thus know, on a scale $[1, 5]$, whether they are high or low in the trait of Extraversion or Neuroticism. Then, to determine whether, say, introverts tend to unfriend people, one needs to know who unfriended whom. Since we do not know that, we characterize each relationship with three values: we take the two personality scores of the individuals in the relationship and compute their minimum, maximum and average values. By running multiple *t-test* on these values, we find that, compared to broken relationships, persistent one tend to have minimum, maximum, and average scores that are higher for Extraversion and lower for Neuroticism (in both cases, the *t-test* have p-values < 0.001), as we have hypothesized.

DISCUSSION

Based on the previous results, we now discuss some open questions.

Data bias. One problem with this kind of studies is to ensure that the sample is not too biased. As we have mentioned previously, validity tests on the personality data suggest that users have responded accurately to the questions of the personality test. Also, our sample does not contain users who are more active than the average one: the median number of contacts for our users is 124, while statistics published by Facebook report an average of 130 for the general population. Similar representativeness holds for gender and age distributions - they are both representative of the wider Facebook

population.

Practical Implications. In addition to advancing current studies on decay of social relations, our findings might well inform the design of a variety of social-networking tools. To see why, consider that we have shown that the propensity of a tie to disappear increases between two individuals who: are not embedded in the same social circles; have no common female friends; have large age difference; and are neurotic or introvert. Since social-networking sites already collect this type of data (especially number of common friends, sex, and age), they could as well offer new monitoring tools, including:

- *People recommender systems.* In Facebook, this would translates into having well-informed ways of populating friend suggestions (i.e., "People you may know" windows).

- *Privacy tools.* One could imagine building tools that inform users about relationships that are about to break or, more likely, tools that fine-tune the broadcasting of Facebook personal updates (e.g., daily updates could be widely broadcasted to people with whom one has persistent relationships and could be selectively shared with people with whom one has relationships that are bound to break).

REFERENCES
1. Bell, R. R. *Worlds of friendships*. Sage Publications, 1981.

2. Costa, P., and Mccrae, R. *The Revised NEO Personality Inventory (NEO-PI-R)*. SAGE Publications, 2005.

3. Gilbert, E., and Karahalios, K. Predicting tie strength with social media. In *Proceeding of the 27th ACM Conference on Human Factors in Computing Systems (CHI)* (2009).

4. Goldberg, L., Johnson, J., Eber, H., Hogan, R., Ashton, M., Cloninger, R., and Gough, H. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality 40* (February 2006).

5. Granovetter, M. Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology 91* (1985).

6. Incite, N. Friends & Frenemies: Why We Add and Remove Facebook Friends. State of Social Media Survey, December 2011.

7. Jong, G. D., Sonderen, E. V., and Emmelkamp, P. A comprehensive model of stress: the roles of experience stress and Neuroticism in explaining the stress- distress relationship. *Psychotherapy and Psychosomatics 68* (1999).

8. Karney, B. R., and Bradbury, T. N. The longitudinal course of marital quality and stability: a review of theory, methods and research. *Psychological Bulletin 118* (July 1995).

9. Kirk, D. M. Gender clustering in friendship networks: some sociological implications. *Methodological Innovations Online 4* (2009).

10. Kivran-Swaine, F., Govindan, P., and Naaman, M. The Impact of Network Structure on Breaking Ties in Online Social Networks: Unfollowing on Twitter. In *Proceeding of the 29th ACM Conference on Human Factors in Computing Systems (CHI)* (2011).

11. Kwak, H., Chun, H., and Moon, S. Fragile OnlineRelationship: A First Look at Unfollow Dynamics in Twitter. In *Proceeding of the 29th ACM Conference on Human Factors in Computing Systems (CHI)* (2011).

12. Martin, J. L., and Yeung, K. T. Persistence of close personal ties over a 12-year period. *Elsevier Social Networks 28* (2006).

13. McPherson, M., Smith-Lovin, L., and Cook, J. M. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology 27* (2001).

14. Nettle, D. *Personality: What Makes You the Way You Are*. Oxford University Press, 2007.

15. Onnela, J. P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J., and Barabási, A. L. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences (PNAS) 104*, 18 (May 2007).

16. Oxford. English Dictionary, 2010.

17. Paul, P. How to Get Unfriended on Facebook. The New York Times, October 2010.

18. Romero, D. M., and Kleinberg, J. The directed closure process in information networks with an analysis of link formation on twitter. . In *Proceeding of the AAAI Conference on Weblogs and Social Media (ICWSM)* (2010).

19. Rubin, L. On men and friendship. *Psychoanalytic Review 73* (1986).

20. Sibona, C., and Walczak, S. Unfriending on Facebook: Friend Request and Online/Offline Behavior Analysis. In *Proceeding of the 44th IEEE International Conference on System Sciences (HICSS)* (2011).