6-2012

More of a receiver than a giver: Why do people unfollow in Twitter?

Haewoon KWAK
Singapore Management University, hkwak@smu.edu.sg

Sue MOON

Wonjae LEE

Citation

KWAK, Haewoon; MOON, Sue; and LEE, Wonjae. More of a receiver than a giver: Why do people unfollow in Twitter?. (2012). *Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, Dublin, Ireland, June 4-7*. 499-502.

Available at: https://ink.library.smu.edu.sg/sis_research/6098

Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media

More of a Receiver Than a Giver: Why Do People Unfollow in Twitter?

Haewoon Kwak* **Sue Moon†** **Wonjae Lee‡**
*Telefonica Research, Barcelona, Spain
†Department of Computer Science, KAIST, Korea
‡Graduate School of Culture Technology, KAIST, Korea

Abstract

We propose a logistic regression model taking into account two analytically different sets of factors–structure and action. The factors include individual, dyadic, and triadic properties between ego and alter whose tie breakup is under consideration. From the fitted model using a large-scale data, we discover 5 structural and 7 actional variables to have significant explanatory power for unfollow. One unique finding from our quantitative analysis is that people appreciate receiving acknowledgements from others even in virtually unilateral communication relationships and are less likely to unfollow them: people are more of a receiver than a giver.

Introduction

Tie creation and breakup are two fundamental processes in the study of social networks and their evolutions. While tie creation has been incorporated as a basic building block in network generative models and social theories, tie breakup has not received equal attention, largely due to lack of data.

Social networks are widely considered interwoven with stable dyadic relationships and finding their way to a state of equilibrium (Friedkin 2004). Two competing theories of action explain why individual actors seek to stabilize their social relationships (Krackhardt 2009). Firstly, individual actors stabilize their social relationships because they adhere to the norm of reciprocity (Gouldner 1960). The social norm is diffuse as to the obliged (Coleman 1990). Individual actors are induced to initiate and maintain a social relationship without certainty of reciprocation. Secondly, the rational choice theory explains the relational stability in terms of exchange equity. Actors cling to their social relationship only if they get even or better off from the exchange (Blau 1986).

The norm of reciprocity does not tell the actors when and who would reciprocate. It is hard to tell whether reciprocation is happenning or not, and thus for actors to stop unfavorable relationships. On the other hand the rational choice theory opens doors to the reasons behind relationship cancellation. If an actor can make a choice based upon exchange equity, it would be an easy decision for him to stop the relationship when it becomes exploitive.

Twitter data offers a unique chance to test whether the actors would stop their social relationships on the basis of exchange equity for two reasons. First, unlike other online social networks (OSNs), we can track down the records of tie breakups (unfollow) over time. More importantly, unlike other OSNs, tie breakup is not against the social norm in Twitter. Twitter users are relatively relieved of the burden of reciprocation and relational stability. A user can follow and unfollow without the others consent, the latter of which is hardly visible to others in the network.

Two recent studies report that unfollow is frequent (Kwak, Chun, and Moon 2011; Kivran-Swaine, Govindan, and Naaman 2011). Kwak *et al.* monitor the changes of follow networks of one million Korean-speaking users during 51 days and find that unfollow is common; 43% of active users unfollow at least once. Kivran-Swaine *et al.* show that structural properties, such as reciprocity and follower overlaps, are associated with the unfollow behavior. We extend these studies by an insight that not only structural properties but also actions between users speak volumes about the status of their relationships. Has a followee sent a reply to a user? Do a followee and a follower share common topics of interest? These actional properties reflect the actual status of relationships and have direct impact on unfollow.

In this paper we build a logistic regression model taking into account two analytically different sets of factors–structure and action. The factors also reflect individual, dyadic, and triadic properties between ego and alter. Our model ends up with 5 structural and 7 actional factors that have significant explanatory power for the odds of unfollow in Twitter.

Most of the variables map to sociological mechanisms, such as homophily, link exchange, or equivalence. One unique finding from the quantitative analysis is that people appreciate the sign of listening from the other party even in virtually unilateral communication relationships. A user is less likely to unfollow those who have *sent* replies or retweets to him than those he sent them to. This result calls for attention to the action mechanism behind social relationships and networks. Tie breakup is better explained by actors' calculation of exchange equity, whereupon the actors appear to be more self-interested than generally considered

*This work is conducted while Haewoon Kwak is at KAIST.

in OSNs. Actors appreciate receiving (re)tweet more than they give.

Data Collection

Twitter offers no official record of unfollow. Instead, we collect snapshots of follow relationships for sampled users and detect unfollowed relationships by comparing consecutive snapshots. We sample users sharing the same cultural context, for cultural beliefs about relationships affect the unfollow behavior. We chose Korean Twitter users for our familiarity with the language and culture. The Korean user community is reasonably large and growing; Korean is the 7th most used language on Twitter as of 2011. We take two snapshots of 1.2 million users' follow networks of Korean users at t_0 of June 25th, 2010 and t_1 of April 26th, 2011. We focus on users who appear in both t_0 and t_1 so as to detect broken relationships between t_0 and t_1. As a result, the number of remaining users is $700,956$ and their follow relationships are $41,920,812$. We also collect up to $3,200$ tweets per user, the upper bound set in Twitter API, at t_0. To address the interdependence between dyads, we sample one million follow relationships from our Korean Twitter graph; that is, we perform random edge sampling. We experiment with three independent samples and obtain consistent results. Unlike user interviews, our electronic behavioral records avoid interviewer effects (Marsden 2003), inaccuracy in recall (Brewer and Webster 2000), and other errors in measurement (Bernard et al. 1984; Marsden 1990; Feld and Carter 2002).

Candidate Independent Variables

Prior to building a model, we identify candidate variables that affect unfollow. Through brainstorming guided by previous literature in sociology and related fields, we pick 78 candidate variables in the following categories: homophily, link exchange, tie strength, power and social hierarchy, informativeness, and attractiveness. These variables represent individual, dyadic, and triadic properties between ego and alter whose tie breakup is under consideration. We note that all independent variables are extracted from the first snapshot only. It enlarges the applicability of our work, as our model works with a single snapshot of a follow network. In the rest of the paper we denote a user as u (ego) and one of followees as f (alter).

Building a Regression Model

After filtering a high correlation among 78 variables, out of 48 candidate variables, we select 23 variables by a stepwise regression based on the Bayesian information criterion (BIC) (Kabaila 2002). With the winnowed-down 23 variables we start building a logistic regression model first with structural properties and then actional properties. Among structural and actional properties, we add individual, dyadic and triadic properties in turn. We have six models in total. Models 1 to 3 incorporate structural properties and models 4 to 6 encompass actional properties of u, f, and the relationship between u and f on unfollow, respectively. Model 1 is the most simple, while Model 6 is the biggest with the most number of variables. We use R, the statistical computing package, on a server with 256 GB memory for the computation.

	Predictors	Odds ratio	95% CI
STRUCTURAL properties			
u	followees	1.000	[1.000, 1.000]
u	followers	1.000	[1.000, 1.000]
u	followees/followers	0.999^{***}	[0.999, 0.999]
u	follow-back ratio	0.712^{***}	[0.672, 0.754]
f	followees	1.000	[1.000, 1.000]
f	follow-back ratio	5.885^{***}	[5.436, 6.371]
$u{\leftrightarrow}f$	link exchange	0.535^{***}	[0.511, 0.560]
$u{\rightarrow}f$	order of follow	1.000	[1.000, 1.000]
$u{\rightarrow}f$	^{N}order of follow	1.000	[1.000, 1.000]
$u{\rightarrow}x{\leftarrow}f$	^{N}common followees	0.001^{***}	[0.001, 0.002]
ACTIONAL properties			
u	mentioned	1.000	[1.000, 1.001]
u	retweeted	0.999^{***}	[0.999, 0.999]
u	favorited	1.000	[1.000, 1.000]
u	tweets	1.000	[1.000, 1.000]
u	tweets containing URL	1.000	[1.000, 1.000]
f	mentioned	1.000	[0.999, 1.000]
f	retweeted	1.000	[0.999, 1.000]
$u{\rightarrow}f$	favorites	0.955^{***}	[0.940, 0.969]
$u{\leftarrow}f$	retweets	0.951^{***}	[0.926, 0.974]
$u{\leftarrow}f$	^{N}replies	0.071^{***}	[0.034, 0.135]
$u{\leftrightarrow}f$	common hashtags	0.937^{***}	[0.915, 0.960]
$u{\leftrightarrow}f$	^{N}common hashtags	0.004^{***}	[0.001, 0.020]
$u{\rightarrow}f$	days since first comm.	0.999^{***}	[0.998, 0.999]
	* $p < 0.05$	** $p < 0.01$	*** $p < 0.001$

Table 1: Odds ratio (rounded to *thousandths*) and its 95% confidence intervals for each variable in Model 6. Superscript^N represents normalized data [0,1]. CI of odds ratio not including 1.000 are highlighted

Results

Model 6, the biggest model, performs the best confirmed by Akaike information criterion (AIC). Also, the analysis of deviance confirms the significance of larger models (all p-values $< .001$). Thus, we focus only on Model 6 from here on. We omit the the outcome of the regression of six models, including estimated coefficients, the standard errors for the estimated coefficients, and p-values, due to lack of space.

In Table 1 we highlight 12 variables whose confidence intervals do not include 1. The odds ratios of all the highlighted variables but for the follow-back ratio of f is less than 1. It means that the likelihood of unfollow decreases when each of 11 variables increases, whereas that of unfollow increases when the follow-back ratio of f increases. We check that the variance inflation factors (VIF) for all these variables are less than 1.5. It reconfirms that variables are not highly correalted.

Interpretation of the 12 variables

We explicate the 12 variables one by one in the context of relevant sociological literature. We then compare the our results with recent unfollow studies (Kwak, Chun, and Moon

2011; Kivran-Swaine, Govindan, and Naaman 2011) and highlight the contributions of this work.

Structural properties The ratio of the number of followees to that of followers is negatively correlated with the likelihood of unfollow, while its power is marginal. It is likely that people with more followees than followers try to retain existing followees.

The follow-back ratio of u and that of f work in the opposite direction; the higher the follow-back ratio of u is, the less likely to unfollow f, but the higher the follow-back ratio of f is, the more likely u is to unfollow f. It can be interpreted that u pays attention to whether f follows back often or not. When the follow-back ratio of f is low, u appreciates f's follow and is less likely to unfollow f. However, when the follow-back ratio of f is high, f's follow does not deter u's unfollow. This relates with status hierarchies discussed in (Gould 2002); we can interpret that f whose follow-back ratio is higher has lower status, and the relationship from u to f of lower status is weak. In the context of Twitter we need a qualitative analysis to support this view, but we leave it for future work.

A unidirectional follow relationship has $\frac{1}{0.535} = 1.869$ times higher odds of unfollow than bidirectional relationships. Although people have inherent differences in the needs of social interaction (Aukett, Ritchie, and Mill 1988), social interaction "pervades every relation of primitive life" (Thurnwald 1932). Bidirectional follow relationships bring emotional closeness, as 'friends' in Facebook do, and thus decrease the likelihood of being unfollowed. Evolutionary game theory models stress the importance of maintaining links with cooperators by "link exchange" (Rand, Arbesman, and Christakis 2011). Our findings also underline the importance of link exchange from the perspective of relationship retention in Twitter.

The normalized number of common followees, calculated by the Jaccard coefficient, is one of the most important variables in unfollow behavior. If the proportion of common followees to the union of followees between u and f decreases by 10%, the odds of unfollow increases by $\frac{1}{exp(-2.885 \times 0.1)} = 1.334$ times. Since the number of common followees is highly correlated to the numbers of followers, neighbors, and transitivities (we filtered them based on correlation coefficients), we can substitute the common followees with the three other variables above and conclude similarly. Common neighbors mean overlapping social circles. Such people retain strong (social) ties and are less likely to break up (Granovetter 1973). This result agrees with the recent findings from qualitative interviews about the top ten people never to unfollow (Kwak, Chun, and Moon 2011). One more interesting observation is that the common followees here are a normalized number; the absolute number of common followees is less important than the ratio of common followees. This outcome calls for a different interpretation from (Shi, Adamic, and Strauss 2007) where a strong tie is defined as 'possessing more than τ mutual friends'. From our model we conclude the ratio of common followees is more important in relationship retention than the absolute number of common followees.

Actional properties Of seven variables of actional properties retweets and favorites account for three: retweeted, favorites, and retweets. The first can be interpreted as follows: the more u is retweeted by people, the less likely u is to unfollow any. Favorites are commonly used for personal archiving of valuable tweets. It thus is reasonable that there exists a negative correlation between the number of favorites and the likelihood of unfollow. We explain the last as follows: the more f retweets u's tweets, the less likely u is to unfollow f. However, u' retweeting of f's tweets is not correlated to u's unfollow of f.

The ratio of replies from f to u is also important. From above we see that more retweets or replies from f to u decrease the likelihood of unfollow by u, whereas those from u to f do not matter. This indicates that people are likely to retain ties with those who express active signs of subscription rather than with those they themselves pay explicit attention to. Similar phenomena exist across OSNs. Burke *et al.* studied social capital among Facebook users via interviews and found that receiving messages, not sending, is associated with increase in bridging social capital (Burke, Kraut, and Marlow 2011). Our results from quantitative analysis are in agreement with their work, even though Twitter is different from Facebook (Kwak et al. 2010). Even in a phone-call network, the persistence of an edge from i to j depends more on the ratio of j's calls to i (odds ratio = 2.345) than that of i's calls to j (odds ratio = 1.052) (Raeder et al. 2011). Gould said, "when people care sufficiently about symmetry (relative to quality), those who receive few or no attributions from the most desirable person will prefer to direct their own attributions toward less attractive alters who at least return the favor" (Gould 2002). This quote perfectly explains above three observations in Twitter, Facebook, and phone calls.

Homophily as in topical similarity in written tweets is an important factor in unfollow. When the Jaccard coefficient of hashtags used by u and f decreases 10 %, the odds of unfollow by u increases by $\frac{1}{exp(-5.54 \times 0.1)} = 1.740$ times. This is in agreement with previous studies in sociology that homophilous ties tend to persist (Suitor and Keeton 1997), and consistent with a Twitter study (Weng et al. 2010) that showed topical similarity is one of the reasons to establish and retain the follow relationships.

The number of days since the first communication has negative correlation with the likelihood to unfollow. Gilbert and Karahalios showed that tie strength is correlated with days since both the first communication and the last communication in Facebook (Gilbert and Karahalios 2009). We confirm that also in Twitter the duration of a tie enforces retention.

We compare our results with two recent studies of unfollow in Twitter. The study of (Kwak, Chun, and Moon 2011) reports that link exchange, common followees, the order of follow relationship, and the number of retweets and favorites are correlated with unfollow. The outcome from our model supports those findings and explicate an exhaustive list of variables quantitaively. However, our cross-sectional data separated by ten months does not capture short-lived follow relationships in between, and thus, the explanatory power of the temporal order of follow relationships seems to have decreased in our model. The model in (Kivran-Swaine, Govindan, and Naaman 2011) is based only on structural variables and does not account for high correlation among variables. Yet our results are generally in alignment with theirs.

The Twitter users appear to be more self-interested in their exchange relationships. They appreciate receiving more attention than giving, and it is pronounced in the act of unfollow. To test this action mechanism, we regress unfollow upon a multitude of variables characterizing exchange equity theory. The result is consistent with our expectation. When a relationship is profitable in (re)tweet exchange, the actor is less likely to unfollow the exchange partner.

Discussion and Future Work

Although we use only one dataset of a common language and cultural background, our findings are well explained by established theories in sociology. Yet there exist differences in the usage of Twitter across languages (Hong, Convertino, and Chi 2011). Twitter is not much like OSNs (Kwak et al. 2010), but our results mostly match findings from other services (Burke, Kraut, and Marlow 2011; Raeder et al. 2011). As pointed out in (Gilbert 2012), predictive power of a model in one service could generalize to another. For future work we would like to apply our model to datasets from different cultural groups and OSNs and demonstrate generalizability of our model.

Acknowledgments

This work was supported by the IT R&D program of MKE/KCA [Project No: 08-911-05-002, "CASFI: High-Precision Measurement and Analysis Research"].

References

Aukett, R.; Ritchie, J.; and Mill, K. 1988. Gender differences in friendship patterns. Sex Roles 19:57–66. 10.1007/BF00292464.

Bernard, H.; Killworth, P.; Kronenfeld, D.; and Sailer, L. 1984. The problem of informant accuracy: The validity of retrospective data. Annual Review of Anthropology 13:495–517.

Blau, P. M. 1986. Exchange and power in social life. Transaction Publishers.

Brewer, D., and Webster, C. 2000. Forgetting of friends and its effects on measuring friendship networks. Social Networks 21(4):361–373.

Burke, M.; Kraut, R.; and Marlow, C. 2011. Social capital on facebook: differentiating uses and users. In Proc. CHI.

Coleman, J. 1990. Foundations of social theory. Belknap Press of Harvard University Press.

Doreian, P. D. 1981. Estimating linear models with spatially distributed data. Sociological Methodology 12:359–388.

Feld, S., and Carter, W. 2002. Detecting measurement bias in respondent reports of personal networks. Social Networks 24(4):365–383.

Friedkin, N. E. 2004. Social cohesion. Annual Review of Sociology 30:409–425.

Gargiulo, M., and Benassi, M. 2000. Trapped in your own net? Network cohesion, structural holes, and the adaptation of social capital. Organization Science 11(2):183–196.

Gilbert, E., and Karahalios, K. 2009. Predicting tie strength with social media. In Proc. CHI.

Gilbert, E. 2012. Predicting tie strength in a new medium. In Proc. CSCW.

Golder, S. A., and Yardi, S. 2010. Structural predictors of tie formation in twitter: Transitivity and mutuality. In Proc. SocialCom.

Gould, R. V. 1991. Multiple networks and mobilization in the paris commune, 1871. American Sociological Review 56:716–729.

Gould, R. V. 2002. The origins of status hierarchies: A formal theory and empirical test. American Journal of Sociology 107(5):1143–1178.

Gouldner, A. W. 1960. The norm of reciprocity: A preliminary statement. American Sociological Review 25(2):pp. 161–178.

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

Hong, L.; Convertino, G.; and Chi, E. H. 2011. Language matters in twitter: A large scale study. In Proc. ICWSM.

Kabaila, P. 2002. On variable selection in linear regression. Econometric Theory 18(04):913–925.

Kivran-Swaine, F.; Govindan, P.; and Naaman, M. 2011. The impact of network structure on breaking ties in online social networks: Unfollowing on Twitter. In Proc. CHI.

Krackhardt, D. 2009. A plunge into networks. Science 326(5949):47–48.

Kwak, H.; Lee, C.; Park, H.; and Moon, S. 2010. What is Twitter, a social network or a news media? In Proc. WWW.

Kwak, H.; Chun, H.; and Moon, S. 2011. Fragile online relationship: A first look at unfollow dynamics in Twitter. In Proc. CHI.

Marsden, P. 1990. Network data and measurement. Annual Review of Sociology 435–463.

Marsden, P. 2003. Interviewer effects in measuring network size using a single name generator. Social Networks 25(1):1–16.

McPherson, M.; Lovin, L. S.; and Cook, J. M. 2001. Birds of a feather: Homophily in social networks. Annual Review of Sociology 27(1):415–444.

Onnela, J.-P.; Saramāki, J.; Hyvönen, J.; Szabó, G.; Lazer, D.; Kaski, K.; Kertész, J.; and Barabási, A.-L. 2007. Structure and tie strengths in mobile communication networks. PNAS 104(18):7332–7336.

Raeder, T.; Lizardo, O.; Hachen, D.; and Chawla, N. V. 2011. Predictors of short-term decay of cell phone contacts in a large scale communication network. Social Networks 33(4):245 – 257.

Rand, D. G.; Arbesman, S.; and Christakis, N. A. 2011. Dynamic social networks promote cooperation in experiments with humans. PNAS 108(48):19193–19198.

Shi, X.; Adamic, L. A.; and Strauss, M. J. 2007. Networks of strong ties. Physica A 378(1):33–47.

Suh, B.; Hong, L.; Pirolli, P.; and Chi, E. 2010. Want to be retweeted? Large scale analytics on factors impacting retweet in Twitter network. In Proc. SocialCom.

Suitor, J., and Keeton, S. 1997. Once a friend, always a friend? Effects of homophily on women's support networks across a decade. Social Networks 19(1):51–62.

Thurnwald, R. 1932. Economics in Primitive Communities. Oxford University Press.

Wasserman, S., and Faust, K. 1994. Social Network Analysis: Methods and Applications. Cambridge University Press.

Weng, J.; Lim, E.-P.; Jiang, J.; and He, Q. 2010. TwitterRank: finding topic-sensitive influential Twitterers. In Proc. WSDM.