Structures of Broken Ties: Exploring Unfollow Behavior on Twitter

Bo Xu, Yun Huang
Northwestern University
2145 Sheridan RD, TECH C210
Evanston, IL, 60208, U.S.A.
{bo.xu, yun}@northwestern.edu

Haewoon Kwak
Telefonica Research
Diag. 00 Plaza de Ernest Lluch i Martín, Barcelona, Spain
kwak@tid.es

Noshir S. Contractor
Northwestern University
2145 Sheridan RD, TECH D241
Evanston, IL, 60208, U.S.A.
nosh@northwestern.edu

ABSTRACT
This study investigates unfollow behavior in Twitter, i.e. people removing others from their Twitter following lists. Considering the interdependency and dynamics of unfollow decisions, we use actor-oriented modeling (SIENA) to examine the impacts of reciprocity, status, embeddedness, homophily, and informativeness on tie dissolution. Focusing on ordinary users in tightly-knitted user groups, the results show that relational properties play key roles in the emergence of unfollow behavior: mutual following relations and common followees reduce the likelihood of unfollowing. And unfollow tends to be reciprocal: when a user is unfollowed by someone, he or she will unfollow back. However, there is no evidence of the impacts of homophily based on common interests and informativeness of interactions. The findings suggest that Twitter has many heterogeneous user groups and relational and informational factors may not be applicable universally.

Author Keywords
Unfollow relations; tie dissolution; Twitter; actor-oriented modeling (SIENA); snowball sampling.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
As a social networking service Twitter helps people to stay connected and share their status through short text messages [6]. With the rapid development all over the world, it has soon become new media for broadcasting and commercial applications. Although many ordinary users socialize with their friends, family, and co-workers in Twitter, famous figures and companies have adopted the media platform to maintain, improve, and promote their public relations. Thus, relation building and information retrieval becomes the two major motivations that drive people to use Twitter [8].

Previous studies show that both relational and informational factors are important for all users in maintaining following relations in Twitter [10]. However, no research has distinguished the two: whether they have the impact at the same time or they exist in different user groups? To be more specific, do all users use Twitter for both relational and informational purposes or do various types of Twitter users have different motivations? This paper explores the different impacts of relational and informational factors on user relations in Twitter.

In Twitter, users can take two steps to establish and maintain their relations with others. When joining Twitter, users may subscribe to others' tweets, known as *follow*, and become followers of other users. Later on, users are free to unsubscribe and remove others from their following lists, known as *unfollow*. Follow and unfollow decisions are distinct and based on different information and reasons.

While a substantial body of studies has investigated the creation and maintenance of network ties, very little research focuses on the dissolution of network ties [13]. Unfollow on Twitter is therefore an excellent opportunity to help us understand the concerted effort to dissolve relations. Further, we focus on unfollow behavior because we believe unfollow decisions are more rational than follow decisions. Twitter's internal suggestions and the prevalence of social toolbars make follow actions as simple as one-click even without visiting Twitter's website. It is possible to follow many users quickly and therefore the follow decisions could be influenced by many external factors. In contrast, the decision of removing specific following relations is more consistent since it is likely based on previous interactions with the targets. Moreover, users' unfollow behavior indicates that they actually read tweets and manage their relations accordingly. The noises in the data due to inactive users can be reduced by studying unfollow.

Tie formation and dissolution are the two equally important factors that drive the structural evolution of a network but the process of tie dissolution is not well covered in the literature perhaps due to the lack of longitudinal data. Most ties in real life decay gradually so that it is very difficult to define the point of a "formal breakup" [11]. However, online social networking services such as Twitter provide temporal information of online relations. Through multiple snapshots of following relations in Twitter, we can detect the breakup of following ties and therefore study the unfollow behavior.

Kwak et al. [9] study unfollow behavior in a Korean Twitter community and find that reciprocity, informativeness, and the overlap of relationships are crucial in the decision to unfollow. This indicates that Twitter, as a whole, is both relational and informational. Kivran-Swaine et al. [7] analyze the structural properties that correlate with unfollow behavior using logistic regression. They conclude that reciprocity and transitivity is most influential to unfollow behavior. However their model does not incorporate the attributes of Twitter users.

Previous research performs logistic regression analysis on the entire user population and assumes that they all have the same behavioral pattern no matter which type of groups they belong to. This paper aims to provide more insightful ideas about user behavior within different user groups. Individuals' motivations of using Twitter may vary from group to group. Unfollow behavior in tightly knitted communities is not the same as that in superstar fan groups. Unfollowing a friend and unfollowing a celebrity should have different reasons. Thus, it is problematic to use the entire population to analyze heterogeneous users since some of them may be influential outliers and distort the results.

Using logistic regression, previous studies assume that unfollow activities are independent of each other. However, users' decisions are related to others' behavior in Twitter. Some users are likely to be unfollowed by many others for the same reasons such as spam. These unfollow decisions against the same user are correlated because of the influence of the common factor. Another example is reciprocal unfollow: when users are unfollowed by others they are likely to unfollow back in response.

This paper investigates different impacts of relational and informational factors on unfollow behavior in Twitter. We focus on ordinary users in tightly knitted user groups and analyze the dynamics of their unfollow relations using a longitudinal dataset with multiple snapshots of Twitter following relations.

We use SIENA, an actor-oriented model, to address the interdependency among unfollow relations and utilize the longitudinal samples [20]. Compared to logistic regression, SIENA assumes that users represented by nodes in a network play a crucial role in changing their ties to others according to their own attributes, others' attributes, and the relations among them. This method of social network analysis incorporates the interdependency of individual decisions through specific network structures. Moreover, SIENA is a dynamic model which uses multiple snapshots of a network at different time points and therefore captures temporal patterns among users, i.e. the tendency for users' unfollow and then re-follow actions in a period of time.

THEORIES AND HYPOTHESES

Different user groups are established based on different types of social bonds. Wittel [23] argues that "community entails stability, coherence, embeddedness and belonging. It involves strong and long-lasting ties, proximity and a common history or narrative of the collective." Communities consist of narrational relations, i.e. relation-oriented and more durable social bonds. However, if the relations in a network are primarily informational and based on ephemeral exchange of data, this network is considered as expressing network sociality rather than a community. Most computer-based virtual communities, working groups, and new media are typical examples of network sociality in which social bonds are created on an informational basis [23].

Previous studies on Twitter show that *reciprocity* [16], *social status* [2], *embeddedness* [21], *homophily* [12] and *informativeness* [8] play key roles in the process of tie formation. This reveals the social network and news media perspectives of Twitter [8]. However, it is not clear whether social networks and media are two *sides* of Twitter or two *parts* of Twitter. In other words, do users use Twitter for both relational and informational purposes at the same time or do various groups of Twitter users have different motivations?

The following paragraphs review five theories of relation building: reciprocity, social status, and embeddedness as relational motivations, and topic-homophily and informativeness as informational motivations. We adopt this set of theoretical frameworks of tie formation to characterize tie dissolution on Twitter. These five factors of relation building may not be applicable for all user groups in Twitter. In tightly knitted user groups studied in this paper, following relations reflect users' social ties and relational factors should play a dominant role in maintaining following ties.

Reciprocity

Reciprocity means behaving towards someone in the manner in which they behave toward you. Paying back a favor and returning a smile with a smile are examples of reciprocal behavior [16]. Previous studies show that reciprocity, as a source of social cohesion, creates stronger mutual ties and increases the stability and equilibrium of the society [3, 18]. In Twitter, mutually following relations enable two users to follow each other's updates, which increase their communication frequency and lead to more stable relations. From this perspective, we have:

Hypothesis 1: Users are less likely to unfollow those who follow them.

On the other hand, reciprocal exchange can also happen in negative interactions. In reciprocal relations, the breaking of one tie indicates decreasing social cohesion between the two. Therefore, it is very likely that the other party will unfollow back in return:

Hypothesis 2: Users are more likely to unfollow those who unfollow them.

Social Status

Status shows the influence and popularity of an individual in a community. According to preferential attachment theory [1], people tend to be better integrated with a social network by connecting to popular nodes in it.

In Twitter, users with many followers are often considered more "powerful." For example, some celebrities have millions of followers and their tweets diffuse much faster and wider in the network. Users with a higher status tend to spend more time on Twitter in order to retain their followers. Therefore other users are more likely to maintain the relations with them.

Hypothesis 3.1: Users with more followers are less likely to be unfollowed.

At the same time, high-status users usually are more serious in managing their relations and followees. They may have a higher tendency to terminate unnecessary relations.

Hypothesis 3.2: Users with more followers are more likely to unfollow.

Embeddedness

Embeddedness is the degree to which individuals are enmeshed in a social network. According to Granovetter [4], the structural embeddedness of a tie between two individuals can be measured by the overlap of social ties between them. In the context of collaboration networks, Newman et al. [14] show that more common neighbors lead to a higher probability of collaboration. Onnela et al. [15] also find a positive correlation between tie strength and the number of common friends.

Since more than 90% of following relations in our sample are mutual, the number of common followees can be used as an approximation of common friends. Therefore the users with more common followees are more embedded in the network.

Hypothesis 4: Users are less likely to unfollow those with whom they share more common followees.

Topic-Homophily

Homophily is the tendency of people to associate with similar people rather than dissimilar ones [11, 12]. By doing so, they avoid potential areas of conflict in the relationship [17].

Weng et al. [22] discover that Twitter users are more likely to follow those who are similar in the topics of their tweets. Kwak et al. [8] report that many respondents in their survey state that they unfollow people because they are not interested in the topics of their tweets, irrespective of the quality.

Many Twitter users include hashtags, that is, the keywords or topics prefixed by the # symbol, in their tweets to label and categorize tweets. In this paper, we use the number of common hashtags as a measure of topic-homophily between users. The more hashtags the two users share in common, the more similarity between them. If the motivation of Twitter following is informational, people are more likely to maintain the relations with similar others for data exchange.

Hypothesis 5: Users are less likely to unfollow those who are interested in similar topics.

Informativeness

Kwak et al. [8] explain the nature of interactions between Twitter users from an informativeness perspective. Informativeness can be measured by the frequency of interactions and data exchange between users in Twitter. Users express their interest in others through interactions such as reply, retweet, mention, and favorite. If people use Twitter relations for information exchange, they are more likely to keep following those who they interacted with.

Hypothesis 6: Users are less likely to unfollow those whom they have retweeted, mentioned, replied, or favorited.

DATA DESCRIPTION

Since cultural beliefs about relationships may vary, we focus on a set of users in the same cultural context. Based on a sample population of *697,628* Korean Twitter users, we took four snapshots of their following relations on June 25th 2010, September 3rd 2010, April 26th 2011, and September 19th 2011. We consider the dissolution of a tie between snapshots as an unfollow relation: a directed link from user A to user B if A followed B but stopped following at the next time point. In this way, the dissolution of following relations is modeled by the emergence of unfollow relations in an unfollow network. For example, there were almost seven hundred thousand users and 34 million following relations on June 25th, 2010, 73.6% of which are mutual. On April 26th, 2011, 858,702 following relations disappeared are used to construct an unfollow network. By comparing the four snapshots, we construct three unfollow networks at different time points: time 1 on September 3rd, 2010, time 2 on April 26th, 2011, and time 3 on September 19th, 2011.

The degree distribution of the first snapshot of following relations is illustrated in Figure 1. The picture shows that the nodes with degree 2,000 or less follow the power law distribution similar to the findings in Java et al. [6]. Users with more than 2,000 followers are mostly superstars and commercial users. Since the aim of this paper is to study user behavior in ordinary user groups, we exclude these users from our analysis.

Due to the heterogeneous structures in Twitter networks, we use one-wave snowball sampling to extract closely connected user groups [5]. The sampling method includes two steps: selecting a sample population based on a seed user and identifying unfollow relations for the sample. First we randomly select a user with 1,000 to 2,000 followers as the seed and find all his or her followers in the first snapshot. Based on their following relations, some outliers with extreme degrees are removed (about 2.5%). And then the unfollow relations among these users are extracted from the three unfollow networks. This approach decomposes the huge Twitter network into many small communities and focuses on the unfollow activities at a normalized level.

We generate 104 random snowball samples. Each sample is estimated independently using SIENA models and all results are combined in a macro-level between-community analysis by using meta-analysis [19].

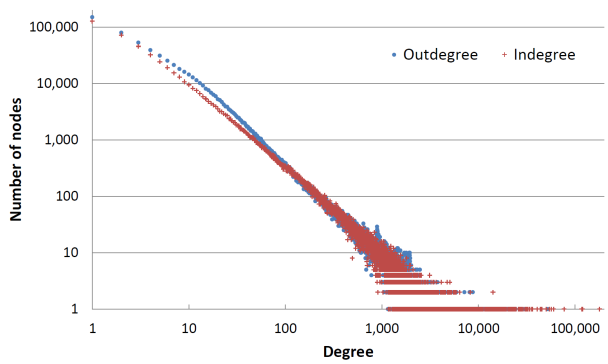

Figure 1. Indegree and outdegree distributions of follow networks on June 25th 2010.

In SIENA models, each user (i.e. sender) makes a decision to form or remove an unfollow relation to another user (i.e. receiver) at each time point based on sender's attributes, receiver's attributes, and the relations between them. Four network statistics are included as control variables to handle the interdependency among the unfollow relations.

MODEL AND RESULTS

As in logistic regressions, SIENA models characterize the impact of explanatory variables on the log odds of links. A positive coefficient in estimation indicates that a larger-valued corresponding explanatory variable leads to a higher tie probability, conditional on all other effects in the model. Table 1 summarizes our model results of the meta-analysis.

The results show that the reciprocity plays a critical role both based on following relations and unfollow relations. Mutual following ties have a negative and significant impact on unfollow relations. When two users follow each other mutually, the odds ratio of one unfollowing another is only 0.63, i.e. exp(-0.46), of those without mutual following relations. Mutual ties indeed make relations stronger and more cohesive. However, if one user in mutual following relations unfollows the other, the unfollowed user is very likely to unfollow back in return. The odds ratio of reciprocal unfollowing is 10.49 times, i.e. exp(2.81-0.46), of that of one-way unfollowing. Both Hypotheses 1 and 2 are supported.

As with following, unfollowing is directed with a sender who initiates the action and a receiver who is unfollowed. Receiver's number of followers has a negative impact on unfollow relations and sender's number of followers has a positive impact. High status users, who have more followers, are less likely to be unfollowed but more likely to unfollow others. Both Hypotheses 3.1 and 3.2 are supported.

Parameters	Estimate (S.E.)	
Relational factors:		
Mutual following ties	-0.46*(.038)	H1: Supported
Reciprocal unfollow	2.81*(.12)	H2: Supported
# Followers (receiver)	-0.05*(.004)	H3.1:Supported
# Followers (sender)	0.09*(.013)	H3.2:Supported
# Common followees	-0.13*(.049)	H4:Supported
Informational factors:		
# Common Hashtag	-0.01 (.008)	H5:Not supported
# Replies	0.004 (.002)	H6:Not supported
# Retweets	-0.009 (.005)	H6:Not supported
# Mentions	0.02 (.009)	H6:Not supported
# Favorites	0.01 (.006）	H6:Not supported
Unfollow network structures as control variables:		
Rate (time 1 to time 2)	32.55*(3.855)	
Rate (time 2 to time 3)	12.93*(1.290)	
Density	-1.97*(.234)	
In-star effect	0.31*(.023)	

Note: * indicates $p<0.05$, # followers and # common followees are in thousands.

Table 1. Summary of model results.

As predicted in Hypothesis 4, the number of *common followees* has a negative impact on unfollow relations. People following many common users are less likely to unfollow each other.

Contrary to Kwak et al. [8], the number of *common hashtags* has no significant impact. There is no evidence that the common interests reduce the likelihood of unfollowing. The topic-homophily effect proposed in Hypothesis 5 is not supported. Similarly, we do not find any significant impacts of *replies*, *retweets*, *mentions,* and *favorites* on unfollow behavior. Hypotheses 5 and 6 are not supported.

Other than the ten explanatory variables drawn from the hypotheses discussed above, *Rate 1*, *Rate 2*, *Density,* and *In-stars effect* are also included to control for the network structures during the Monte Carlo Markov Chain Simulation in SIENA [20]. Compared to logistic regression models, SIENA models consider important network structures when examining the impacts of explanatory variables. *Rate 1* (*Rate 2)* captures the base-line change rates of unfollow relations between time 1 and time 2 (time 2 and time 3). The positive Rate 1 indicates that more unfollow relations are observed at time 2 compared to time 1. *Density*, like the constant (or intercept) in logistic regressions, controls for the number of edges in the networks. The negative coefficient of *density* suggests that the unfollow networks are sparse and people are less likely to unfollow others randomly. In fact only 2.5% of following relations have dissolved in ten months. *In-star effect* controls for the in-degree distribution. The positive coefficient of *in-stars effect* indicates that users are more likely to be unfollowed by many others at the same time. This result demonstrates the interdependency among the unfollow relations.

DISCUSSION

Consistent with previous empirical research, our analysis indicates that relational factors have strong influences on unfollow decisions in Twitter. As predicted by Hypotheses 1-4, reciprocity, status, and embeddedness are key properties that affect user behavior of maintaining following relations. Twitter users seek cohesion and equilibrium when managing their relations and tend to keep reciprocal relations. Only 0.26% of mutual following ties disappeared within 10 months in our samples. This suggests that mutually following friends are unlikely to unfollow each other. High status users with more followers are more serious and active in posting tweets and organizing their following relations. Thus, they are less likely to be unfollowed and more likely to unfollow others. Users connected by many common friends (e.g. followees) are embedded in a strong network structure. These entangled social bonds contribute to the persistence of their relations.

However, the impact of information exchange is inconclusive. Hypotheses 5 and 6 are not supported and topic-homophily and informativeness has no significant impact on unfollow in our sample groups. Although the results contradict previous studies [10], this study demonstrates that by focusing on some parts of the Twitter network, the influences of relational and informational factors can be disentangles.

We believe that Twitter has many heterogeneous user groups and it is not appropriate to model different types of behavior using the entire population as a whole. Using the samples of tightly connected user groups, we show that the relations in these groups are more likely to reflect users' pre-existing social bonds in reality, which represent long lasting, narrative relations. Therefore, informational factors such as the content and frequency of their interactions may not affect the persistence of their following ties.

The measurement issue could be another potential reason for the lack of informational motivations in these groups. Since we did not perform semantic analysis on tweets, the measures used may not fully capture the informational factors of user behavior. For example, common hash tags indicate the overlapping of keywords users tweeted and retweeted but may not truly reveal their shared interests. Similarly, the numbers of replies, retweets, mentions, and favorites measure the frequency of interactions but not the information exchanged.

Our findings also suggest that users make decisions according to other's behavior in Twitter. Previous studies consider each unfollow action as independent and use logistic regressions to model the decisions. Using SIENA models, we show that unfollow decisions are correlated and the network structures have critical impacts: unfollow decisions are highly reciprocal and clustered. For example, the breakup of one tie in a pair of mutual following relations will lead to the breakup of the other. 8% of unfollow relations are in pairs (i.e. reciprocal to each other), which is significantly higher than random chances. Ignoring the interdependency of relations in the Twitter network may lead to wrong conclusions.

CONCLUSION

This paper studies the dissolution of following relations in Twitter. Based on the findings in tightly connected user groups, we show that only relational motivations have significant impacts on maintaining following relations but the impact of topic-homophily and informativeness is not significant. This suggests that relational and informational factors may not influence all users at the same time. It is likely that Twitter has many types of sub-groups with different motivations.

We use small and tightly connected user groups to explore the differences of relational and informational factors in tie dissolution and find relation-oriented groups in Twitter. For other types of users, such as followers of celebrities and interest groups, information oriented motivations such as common interests may have a dominant impact on unfollow behavior. Limited by the sampling methods and focal population, this study cannot reveal the reason why individual interactions such as replies, retweets, mentions, and favorites have no impact on maintaining following ties. Future research will evaluate and compare unfollow patterns in Twitter user groups of different types and sizes. The snowball sampling and meta-analysis used in this paper may provide an adequate approach to study the heterogeneous groups in a large network.

ACKNOWLEDGMENTS

This study was supported by the National Natural Science Foundation (Grant No. CNS-1010904, OCI-0904356, & IIS-0841583) and Army Research Lab (W911NF-09-02-0053), the National Natural Science Foundation of China (No. 90924020), MOE project of Humanities and Social Sciences (No. 11YJA630044), and the Innovation Foundation of BUAA for PhD Graduates. We thank Wonjae Lee and Sue Moon for preparing data sets and CSCW reviewers for providing helpful comments on previous versions.

REFERENCES

1. Barabasi, A. L., Albert, R., Emergence of Scaling in Random Networks, *Science* 286 (5439) (1999) 509-512.
2. Emerson, R. M., Power-Dependence Relations, *American Sociological Review 27* (1) (1962) 31-41.

3. Gouldner, A. W., The Norm of Reciprocity: A Preliminary Statement, *American Sociological Review* 25 (2) (1960) 161-178.
4. Granovetter, M., Economic Action and Social Structure: The Problem of Embeddedness, *Readings in Economic Sociology,* Blackwell Publishers Ltd, 2008, pp. 63-68.
5. Gruzd, A., Wellman, B., Takhteyev, Y., Imagining Twitter as an Imagined Community, *American Behavioral Scientist* 55 (10) (2011) 1294-1318.
6. Java, A., Song, X., Finin, T., Tseng, B., Why we twitter: understanding microblogging usage and communities, *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis*, 2007.
7. Kivran-Swaine, F., Govindan, P., Naaman, M., The impact of network structure on breaking ties in online social networks: unfollowing on twitter, in *Proc. CHI 2011*, ACM Press (2011), 1101-1104.
8. Kwak, H., Chun, H., Park, H., Moon, S., What is twitter, a social network or a news media, *in Proc. www 2010*, ACM Press(2010), 591-600.
9. Kwak, H., Chun, H., Moon, S., Fragile Online Relationship: A First Look at Unfollow Dynamics in Twitter, in *Proc. CHI 2011*, ACM Press (2011), 1091-1100.
10. Kwak, H., Moon, S., Lee, W., More of a receiver than a giver: Why do people unfollow in Twitter, in *Proc. ICWSM 2012*, AAAI Publications (2012).
11. Martin, J. L., Yeung, K.-T., Persistence of close personal ties over a 12-year period, *Social Networks* 28 (2006) 331-362.
12. McPherson, M., Smith-Lovin, L., Cook, J. M., Birds of a Feather: Homophily in Social Networks, *Annual Review of Sociology* 27 (2001) 415-444.
13. Monge, P., Contractor, N., *Theories of Communication Networks*, Oxford University Press, Oxford, UK, 2003.
14. Newman, M. E. J., Clustering and preferential attachment in growing networks, *Physical Review E* 64 (2001) 025102.
15. Onnela, J. P., Saramaki, J., Hyvonen, J., Szabo, G., Lazer, D., Kaski, K., Kertesz, J., Barabasi, A. L., Structure and tie strengths in mobile communication networks, in *PNAS* (2007), 7332-7336.
16. S.A., Golder., Yardi, S., Structural Predictors of Tie Formation in Twitter: Transitivity and Mutuality, in *IEEE Social Computing* (2010), 88-95.
17. Sherif, M., Superordinate Goals in the Reduction of Intergroup Conflict, *American Journal of Sociology* 63 (4) (1958) 349-356.
18. Simmel, Georg, *On individuality and social forms: Selected writings*, University of Chicago Press, 1971.
19. Snijders, T. A. B., Baerveldt, C., A multilevel network study of the effects of delinquent behavior on friendship evolution, *The Journal of Mathematical Sociology* 27 (2-3) (2003) 123-151.
20. Snijders, T. A. B., van de Bunt, G. G., Steglich, C. E. G., Introduction to stochastic actor-based models for network dynamics, *Social Networks* 32 (1) (2010) 44-60.
21. Wasserman, S. a. F., K., *Social network analysis: Methods and applicaitons*, Cambridge University Press, 1994.
22. Weng, J., Lim, E.-P., Jiang, J., He, Q., TwitterRank: finding topic-sensitive influential twitterers, in *Proc. WSDM 2010*, ACM Press (2010), 261-270.
23. Wittel, A., Toward a Network Sociality, *Theory, Culture & Society* 18 (6) (2001) 51-76.