

# Conversation Practices and Network Structure in Twitter\*

**Luca Rossi**

Department of Communication Studies  
University of Urbino *Carlo Bo*  
luca.rossi@uniurb.it

**Matteo Magnani**

Department of Computer Science  
Aarhus University  
magnanim@cs.au.dk

## Abstract

The *public by default* nature of Twitter messages, together with the adoption of the #hashtag convention led, in few years, to the creation of a digital space able to host worldwide conversation on almost every kind of topic. From major TV shows to Natural disasters there is no contemporary event that does not have its own #hashtag to gather together the ongoing Twitter conversation. These topical discussions take place outside of the Twitter network made of followers and friends. Nevertheless this topical network is where many of the most studied phenomena take place. Therefore Twitter based communication exists on two almost autonomous levels: the Twitter network made of followers and friends that shows a certain level of stability and the topical network, characterized by a high level of contingency, that appears and disappears following the rhythm of a worldwide conversation.

Despite the fact that this double nature of Twitter is widely recognized among scholars there is still little literature facing the relationships between these two networks. This paper presents the results of an empirical research aimed at discovering how the Twitter network is affected by what happens on the topical network. Does the participation in the same hashtag based conversation change the follower list of the participants? Is it possible to point out specific social behaviors that would produce a major gain of followers? Our conclusions are based on real data concerning the popular TV show *Xfactor*, that largely used Twitter as the official backchannel platform for its audience.

## Introduction

Twitter is undoubtedly one of the most successful social media platforms in the World. With a user base of more than 200 million users (Bruns 2011) its global success can be compared only to Facebook. Its *public by default* nature transformed the service into a digital place able to host, at the same time, communication addressed to a known list of followers (Marwick and boyd 2010) and global conversations with an undefined audience (Bruns and Burgess 2011).

\*This work has been partly funded by FIRB project *Information monitoring, propagation analysis and community detection in social network sites* and PRIN project *Relazioni sociali ed identità in Rete: vissuti e narrazioni degli italiani nei siti di social network*  
Copyright © 2012, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

This *double nature* of Twitter makes it different from many others SNSs both in respect to the function that it has for the users and in respect to the emergent social phenomena that it is able to host. From a network study perspective this double nature can be clearly observed in Twitter's networks. The network made of all the relationships between the users<sup>1</sup> represents only a small part of all the communication opportunities. On one side it is always possible for a user to address a message to another user by simply adding the @ symbol in front of his/her username; on the other side the messages aggregated through the *hashtags* (Bruns and Burgess 2011) are publicly available and these conversations do not represent any pre-existing network and therefore are usually studied by observing the ephemeral networks created during the conversations (Rossi, Magnani, and Iadarola 2011).

Despite the large interest shown by researchers on Twitter based phenomena there is surprisingly few research that focuses on the relationships between the structural network and what happens during a hashtag based conversation. This paper will move a first step into that direction by investigating how a Twitter conversation can modify the structure of friends and followers of the involved users.

This paper presents two main contributions to the state of the art regarding the analysis of Twitter and SNS data. Beside the practical results obtained by analyzing real data and confirming our experimental hypotheses regarding the interactions between the follower and topical networks and the role of the topical networks in the dynamics of follower acquisition, this paper also introduces an important innovation from a methodological point of view. In particular we show how to apply recent theories about the co-existence of parallel social networks (Magnani and Rossi 2011) to the practical analysis of real social network data. To the best of our knowledge we present the first analysis regarding the interaction between different and parallel networks appearing on Twitter: a more stable network of contacts and a set of highly dynamic and pulsing topical networks. In particular our data have been collected by monitoring a specific hashtag (#XF5) that was used to tweet about the fifth edition of

<sup>1</sup>We use *followers* to identify all the users that follow a specific user and *friends* to identify all the users followed by a specific user (Huberman, Romero, and Wu 2008)

the popular TV show XFactor Italia.

The paper is structured as follow: in the next section we review the existing literature on Twitter-based research and network evolution of SNSs. Then we describe our experimental setting and the major results of our research. Finally we connect these results with the broader debate on SNSs and on Twitter and present our concluding remarks.

### Retaled works

Since its early days the worldwide success of Twitter, which was launched in 2006, has generated a large volume of research. Many of these early studies were focused on the analysis of the network structure and of its topological characteristics (Java et al. 2007). Lately it has become apparent that these researches should be complemented with specific analyses of *local* phenomena that do not necessarily appear at a higher level but still constitute the essence of the Twitter communication network from the point of view of its users. Therefore, beside the attempt to describe the topological aspects of the Twitter network many researchers started to focus on the communicative practices of the platform (Huberman, Romero, and Wu 2008; Marwick and boyd 2010)

The wide diffusion of Twitter has led to the emergence of unexpected social phenomena that found on this SNS a perfect socio-technical environment able to host them. As a consequence of that in few years Twitter has become a digital space where public issues can be discussed, critical information can be shared during natural disasters and TV shows can be commented by and with their fans. More recent researches then focused on these phenomena with a closer perspective (Bruns 2011; Magnani, Montesi, and Rossi 2010a; Rossi, Magnani, and Iadarola 2011). Nevertheless these researches do not focus on the evolution of the structural properties of the network as a consequence of hashtag conversations.

This aspect can also be understood as an issue related to the evolution of a network or, more specifically of evolution of a SNS. This field has seen in recent times a growth of interest specifically related to the analysis of SNSs. Many of these researches explored the relationship between the homophily-driven creation of links and the topological properties of the network (Crandall et al. 2008; Viswanath et al. 2009). The homophily driven approach suggests that users with similar interests will have a higher probability of being connected than users with different interests. Recently, the dynamics of *unfollowing* have also been studied (Kwak, Chun, and Moon 2011). Within this paper we aim at expanding the boundaries of this research line with a study that complements homophily-driven link creation theories.

The theoretical basis of our research consists in the recent literature on complex social networks, that has dealt with several kinds of networks able to represent, within the same graph structure, different kinds of relationships between users (Kazienko, Musial, and Kajdanowicz 2011; Kazienko et al. 2010) or even relationships between users with different attributes (Zhao et al. 2011). These approaches deal with the heterogeneity of ties and also with the multiplicity of user identities (Magnani and Rossi 2011) and they point

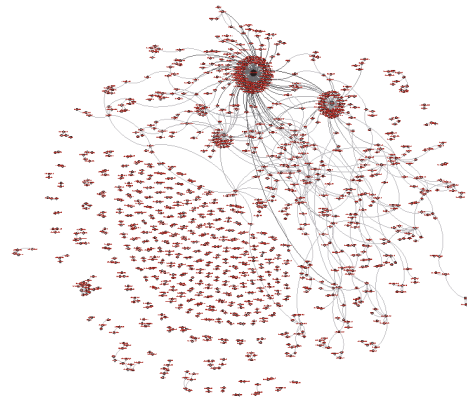


Figure 1: Visualization of the reply messages among the users (image produced using Gephi)

out several crucial aspects that have to be taken into careful consideration while we approach the description of large real-world SNSs.

### Experimental results

In this section we experimentally evaluate the impact of participating to a topical network (and the specific involvement in this network) on the network of followers. We first describe how we acquired the data and present basic statistics, then we analyze the communication patterns in the topical networks, and finally we measure the correlation of these aspects with the dynamics of follower acquisition.

#### Data collection and description

To evaluate the relationships between topical networks and the structural network of followers we have extracted in real time the tweets containing the hashtag #XF5 for all the duration of the TV show, for approximately 6 hours (from 19:28:16 pm of January 5th to 01:31:55 am of January 6th 2012). We also monitored the network of followers of a large sample of the participating users (50%, corresponding to the users who started twitting earlier) once every 20 minutes.

The collected dataset is composed by 22 287 tweets produced at a rate of 3 700 tweets per hour (about one tweet per second). The messages in our dataset have been posted online by 5 137 distinct users – every user posted on average 4.34 tweets with a standard deviation of 7.58. Among these messages 5 382 (24.25%) were retweeting a previous message, 1 214 (5.45%) were replying to a user or to a message and 1 347 (6.04%) contained a link to an external webpage.

#### Communication pattern analysis

A clear understanding of the communicative practices that occurred during the TV show can be obtained by looking at Figures 1 and 2. Figure 1 shows the structure of the reply network existing between the users taking part in the hashtag conversation. Each edge of the network represents a tweet addressed to a specific user through the @ reply system. Size and colour of the nodes are defined according to the



Figure 2: Visualization of retweet messages (Gephi)

degree centrality value. What can be clearly noticed is that the most addressed users, visualised in the upper part of the figure, are three accounts officially related to the TV show: *Xfactor\_italia* (the official Twitter account of the show), *alecattelan* (the twitter account of the presenting TV host) and *Simo\_Ventura* (the official account of one of the judges of the game).

Instead of being a mark of real conversational practices this network shows the Twitter practice of commenting by addressing the comment to a specific user.

Within this perspective it appears reasonable that those accounts who were directly involved in the TV show received the highest number of messages and comments. At the same time it emerges a peculiar inter-media conversation where people at home react to what is happening on TV and have a chance not only to comment on it publicly but also to address their comments directly to the performers.

On the other side Figure 2 shows the retweet chains of the messages. It is worth noticing that even if the official *Xfactor\_italia* account still has a dominant role within the network, the clear quantitative distinction between *accounts officially related to the show* and *audience's accounts* seems to fade. Among the top three accounts with the highest number of retweets received we of course find *Xfactor\_italia* but also two users from the audience (*Titofaraci* and *CloCecchetto*) that are not officially involved in the show.

The differences in the reply net and in the retweet net suggest a different use of the social practices that are available within the Twitter system. While the comment/conversation phase seems to focus on the users' identity, emphasizing those accounts that were directly involved in the TV show, the retweeting phase seems to be more focused on the produced content instead of the producer.

### New followers and Network evolution

While many of these results are largely aligned with previous studies made on Twitter as a backchannel tool for TV audiences, these data will be here applied to verify the main

hypothesis of this paper, and are thus particularly useful now that we want to investigate if these communicative activities have consequences on the network structure in terms of new followers acquisition. Since Twitter does not require any authorization in order to follow a public account we could assume that those users who had been able to gain enough visibility during the event and that produced interesting contents would gain some new follower. Our sample of 2 505 users gained, at the end of the monitoring time, 6 526 new followers (avg. 7.03, std. dev. 51.02). As expected the distribution of new followers is very skewed with VIP users that gained a very large number of followers and a long list of users gaining a few followers. This kind of evolution can be explained by classical fitness models (Barabási 2003) where the most fitted nodes become hubs and attract preferably new links. Within this perspective we could ask if it is possible to detect, from its communicative behaviors, the level of *fitness* of a given Twitter user. Is it possible to observe any kind of correlation between users' behavior during the hashtag conversation and the amount of followers that every user gained?

In order to answer this question we have compared the number of new followers gained by every user with three values able to measure the activity that they had during the hashtag conversation: number of posted tweets, degree centrality on the reply network and degree centrality on the retweet network. Results are summarized in Table 1.

Table 1: Correlation between number of new followers and a) number of tweets, b) degree centrality on the reply network, c) degree centrality on the retweet network

	# Tweets	@-centrality	RT-centrality
Correlation	0.15	0.75	0.61

The data clearly show that while the number of tweets has a very limited relation with the acquisition of new followers, being in a central position in the retweet or in the reply *topical networks* affects in a positive way the chance of gaining new followers *in the structural network*. New followers acquisition has a correlation of 0.61 with the centrality value on the retweet network and a correlation of 0.75 with the centrality on the reply network. As we stated at the beginning of the paper we assume that the participation in the same hashtag based conversation can be taken as an element of homophily between the participants. Therefore we can consider these results even more significant because they show such an evident difference with respect to a homogeneous situation. The sharing of a common interest is nevertheless an important aspect in order to study link creation dynamics but it has to be observed within a more complex frame that takes into consideration also the ephemeral communicative relations that users establish.

### Discussion

During a large hashtag conversation users are usually exposed to a very large number of messages. Even in our dataset, which is quite small if compared to the number of messages that large international media events can generate,



the rate of one tweet per second made almost impossible for the average users to read through every message. Concurrent messages, as already observed in the literature (Magnani, Montesi, and Rossi 2010b; 2011), are in fact a real limit to the chance that every message has to spread and be visible through the network. While users writing many messages surely have a larger opportunity of getting visible, visibility does not seem to be a key element in the establishment of new follower relationships even within a homophily context such as the participation in the same hashtag conversation.

The analysis of users' centrality within the reply network and within the retweet network allowed us to observe the impact of these two social dynamics within the process of link establishment. While surely being widely retweeted or being widely addressed increases the visibility of the user as well as posting a high number of messages, these aspects have a positive correlation with the acquisition of new followers.

This phenomenon could be explained by noticing that the acts of retweeting or mentioning seem to work as actions that pre-select relevant and valuable content. By choosing to retweet or to reply to a specific message Twitter users make some kind of implicit endorsement of that message and of that user. Within a flow of a large number of messages this acts as a social way to select relevant users more efficiently than the simple posting activity. It should be close to everyone's experience that the relevance of speakers has no or few relationship with his or her loquacity. Within the social context of Twitter conversation users seem to use implicit social clues (such as the fact that a specific account is mentioned or retweeted) in order to decide who to follow.

## Conclusion

At the beginning of this paper we rose two main research questions. On one side we asked if the participation to highly social Twitter phenomena such as the hashtag based conversations could have a direct impact on the Twitter network structure made of non reciprocal following relationships. On the other side we asked what kind of social behavior could influence the creation of new links. Is the simple participation to a shared conversation about the same topic enough?

Our data show some interesting perspectives on these problems. Focusing on the first question the answer is yes: being an active author on a widely discussed topic seems to increase the chances of getting new followers. While of course the distribution of new followers is rather skewed every user in our sample obtained on average more than 7 new followers after being active on the #XF5 hashtag. The second question is undoubtedly more complex. While the relationship between the participation to a highly visible Twitter conversation and the acquisition of new followers is intuitive to understand we cannot say the same of the analysis of what kind of social behaviors are correlated with the acquisition of new followers. In fact the simple production of a high number of messages is just weakly correlated with this aspect. In order to have a more direct impact on the number of followers users do not have just to post tweets but those tweets have to produce a social reaction: conversational behaviors such as replies or retweets are strongly related to the acquisition of new followers.

## References

- Barabási, A.-L. 2003. *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*. Plume.
- Bruns, A., and Burgess, J. 2011. #Ausvotes: How twitter covered the 2010 Australian federal election. *Communication, Politics & Culture* 44(2).
- Bruns, A. 2011. How long is a tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society* (2011).
- Crandall, D.; Cosley, D.; Huttenlocher, D.; Kleinberg, J.; and Suri, S. 2008. Feedback effects between similarity and social influence in online communities. In *KDD*. ACM.
- Huberman, B. A.; Romero, D. M.; and Wu, F. 2008. Social Networks that Matter: Twitter Under the Microscope. *SSRN eLibrary*.
- Java, A.; Song, X.; Finin, T.; and Tseng, B. L. 2007. Why we twitter: An analysis of a microblogging community. In *WebKDD/SNA-KDD*.
- Kazienko, P.; Bródka, P.; Musial, K.; and Gaworecki, J. 2010. Multi-layered social network creation based on bibliographic data. In *SocialCom/PASSAT*. IEEE.
- Kazienko, P.; Musial, K.; and Kajdanowicz, T. 2011. Multi-dimensional social network in the social recommender system. *IEEE Transactions on Systems, Man, and Cybernetics, Part A* 41(4).
- Kwak, H.; Chun, H.; and Moon, S. B. 2011. Fragile online relationship: a first look at unfollow dynamics in twitter. In *CHI*. ACM.
- Magnani, M., and Rossi, L. 2011. The ml-model for multi-layer social networks. In *ASONAM*. IEEE.
- Magnani, M.; Montesi, D.; and Rossi, L. 2010a. Friend-feed breaking news: Death of a public figure. In *SocialCom*. IEEE.
- Magnani, M.; Montesi, D.; and Rossi, L. 2010b. Information propagation analysis in a social network site. In *ASONAM*. IEEE.
- Magnani, M.; Montesi, D.; and Rossi, L. 2011. *The Influence of Technology on Social Network Analysis and Mining*. Springer. chapter: Factors enabling information propagation in a Social Network Site.
- Marwick, A. E., and boyd, d. 2010. I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society* 13(1).
- Rossi, L.; Magnani, M.; and Iadarola, B. 2011. #rescatem- ineros, global media events in the microblogging age. In *Selected paper of Internet research*. AOIR.
- Viswanath, B.; Mislove, A.; Cha, M.; and Gummadi, P. K. 2009. On the evolution of user interaction in facebook. In *WOSN*.
- Zhao, P.; Li, X.; Xin, D.; and Han, J. 2011. Graph cube: on warehousing and OLAP multidimensional networks. In *SIGMOD*. ACM.