IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. **1(2)** (2011), pp. 216–220

Structure and Dynamics of Emerging Social Networks from Twitter's Conversation #SOSInternetVE

A.J. Morales[1,†], J.C. Losada[1] and R.M. Benito[1]

[1] *Grupo de Sistemas Complejos, Universidad Politecnica de Madrid*

Abstract. Over the past years, new technologies and specially online social networks have penetrated into world's population at an accelerated pace. In this paper we analyze collected data from the web application Twitter, in order to describe the structure and dynamics of the emergent social networks, based on complexity science. We focused on a Venezuelan protest that took place exclusively by Twitter during December, 2010. We found community structure with highly connected hubs and three different kinds of users that determine the information flow dynamics. We noticed that even though online social networks appear to be a pure social environment, traditional media still holds loads of influence inside the network.

Keywords: Complex Networks, Online Social Networks, Information Flow

† **Corresponding author:** alfredo.moralesg@alumnos.upm.es
Received: October 17, 2011
Published: October 24, 2011

1. Introduction

The fast growth of online social networks among population is changing the way people communicate with each other. This type of communication allows societies to coordinate at an unprecedented scale, accelerating social transformation processes all around the globe. These technologies also provide loads of data that allow research on human behavior [1], [2], [3].

Our main goal is to analyze one of these social manifestations, in order to describe the structure and dynamic that rule these new communication processes.

2. Methodology and Results

We have carried out a quantitative analysis of a virtual protest that took place in Venezuela during December 2010, that reached the most discussed topic in Twitter for over 4 hours. It consisted in posting messages to the web application including the hashtag #SOSInternetVE in the message's text. Hashtags are commonly used to identify messages in Twitter, so people can filter them and find out what others are saying in real time.

Our dataset consists in all messages with the hashtag #SOSInternetVE from the Twitter servers between December 14-19, 2010. It is compound by 421.602 messages, written by 77.706 users. Almost 60% of messages were sent from smart mobile phones. In Fig. 1 we present the evolution of the message rate. It can be noticed that some user activity was recorded after the hashtag's first appearance, but it is between December 16-17, 2010, when it bursts and reaches its highest point, showing critical phenomena features. We found that the distribution of posted messages by user follows a power law as shown in the inset of Fig. 1. This means that most of users only posted less than a couple of messages, while a few of them posted more than 100.

We built two networks relating the users from our dataset, according to different Twitter interactions. On one hand we built a network based on *who follows who*, which means, *who received whose messages*. This network is a subgraph of the global Twitter social network and shows the substratum through which the information may have flowed during the protest. It is a directed and non weighted graph, compound by 77.706 nodes and 5.761.331 links. Both *in* and *out degrees* follow power law distributions as shown in Fig. 2. The mean distance between nodes is $d_F = 2,2$, which is a very low value, due to the effect of highly connected hubs. On the other hand, we built another network according to *who retransmits whose messages*. This network indicates the effective links through which the information actually flows inside the social substratum. It is a network that emerges from the information diffusion dynamics. It is a directed and weighted graph, compound by 54.423 nodes and 231.485 links. Both *in* and *out strength* follow power law distributions as shown in the inset of Fig. 2. The mean distance between nodes is $d_R = 3,4$, which is also very low, due to the presence of hubs.

We found different user behaviors, according to the audience size, retransmission level and messages posted as shown in Fig. 3. It can be noticed that for the same amount of followers a user gets more retransmissions as increases its activity. This result, in addition to the followers degree correlation, shown in Fig. 4, let us classify users into three categories: Information producers, active consumers and passive consumers. The *information producers* are widely followed and gain an enormous amount of retransmissions, whereas they have low activity. They do not tend to follow a lot of people, nor retransmit many messages. These accounts belong to traditional mass media agents like TV, journalists, politicians and celebrities. On the other hand *active consumers* are users with high reciprocity in relations. They tend to gain as much audience and retransmission rate, as the amount of activity employed. They are key in the information diffusion process, because they boost the content and serve as links between sources. At last, *passive consumers* are the largest group of users. They consume more information than what they produce. They are characterized for having low activity rate and receiving messages from more people than their audiences.

We also calculated the community structure for both graphs based on the algorithm described in [4]. On the follower graph, we found six communities grouping more than 98% of the population. By identifying the most followed users, we found that communities are formed around different collectives: opposition media, opposition politicians, entertainment celebrities, international media, comedy accounts and government favorable politicians. Nevertheless, the retransmission graph showed a completely different behavior. In this network we found thirty main communities compound by at least 500 users each. Almost all of the communities contain an information producer highly followed and retransmitted. This result shows that users are more selective when it comes to retransmitting in comparison to just listening.

3. Conclusions

We studied the Venezuelan protest #SOSInternetVE, by constructing two networks that represent the social substratum and the information diffusion graph, respectively. Both networks have degree distributions that follow power laws, and very small mean distances between nodes. We identified three types of user behavior that determine the dynamics of information flow: Information Producers, Active Consumers and Passive Consumers. We also found different

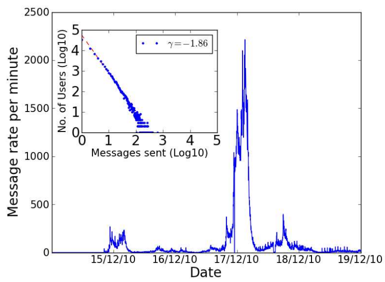

Figure 1: Evolution of message rate. Inset: Distribution of users according to the amount of messages sent.

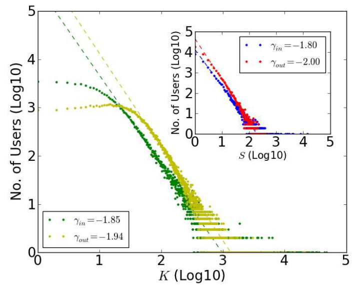

Figure 2: Degrees distribution of the followers network. Yellow and green represent the followers in and out degree. Inset: Strength distribution of retransmission network. Blue and red represents the retransmissions in and out strength.

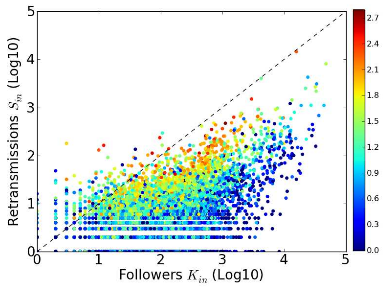

Figure 3: Correlation between followers K_{in} and retransmissions S_{in}, colored by the amount of messages sent.

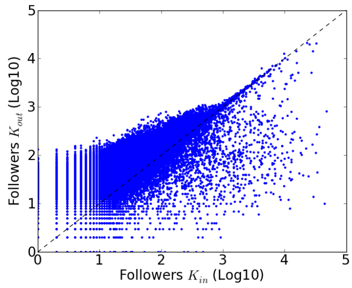

Figure 4: Degrees correlation of the followers network. For each user we represent its in degree vs out degree.

communities around the information producers, and found that people is more selective when it comes to retransmitting information. Finally, we conclude that although the online social media seems to be a pure social phenomena, traditional media agents still enjoy a lot of power and influence over people, who they use to boost and enhance their messages.

Acknowledgements

Support from MICINN- Spain under contracts No. MTM2009- 14621, and i-MATH CSD2006-32, is gratefully acknowledged.

References

[1] David Lazer, et al., *Nature*, 721-723 (Feb. 2009).

[2] H. Kwak, C. Lee, H. Park, and S. Moon, *Proceedings of the 19th international conference on World Wide Web* , (2010).

[3] M. Romero, W. Galuba, S. Asur, and B. A. Huberman., *Technical Report arXiv:1008.1253, CoRR, http://arxiv.org/abs/1008.1253*, Page (2010).

[4] Blondel, V.D. et al., *J. Stat. Mech 10008*, 1-12(2008).