Chaos, Solitons & Fractals 72 (2015) 90–98

Contents lists available at ScienceDirect

Chaos, Solitons & Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Multiple leaders on a multilayer social media

J. Borondo *, A.J. Morales, R.M. Benito, J.C. Losada

Grupo de Sistemas Complejos and Departamento de Física y Mecánica, Universidad Politécnica de Madrid, ETSI Agrónomos, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Available online 28 January 2015

ABSTRACT

Twitter is a social media platform where users can interact in three different ways: following, mentioning, or retweeting. Accordingly, one can define Twitter as a multilayer social network where each layer represents one of the three interaction mechanisms. First, we review the main findings of our previous work regarding two Twitter political conversations: the 2010 Venezuelan protest and the 2011 Spanish general elections. We found that the structure of the follower layer conditions the retweet layer, as having a low number of followers represents a constrain to effectively propagate information. The collapsed directed multiplex network does not present a rich-club ordering, as politicians presided large communities of regular users in the mention layer; while media accounts were the sources from which people retweeted information. However, when considering reciprocal interactions the rich-club ordering emerges, as elite accounts preferentially interacted among themselves and largely ignored the crowd. Finally, we explore the main relationships between the community structure of the three layers. At the follower level users cluster in large and dense communities holding various hubs, that break into smaller and more segregated ones in the mention and retweet layers. Hence, we argue that to fully understand Twitter we have to analyze it as a multilayer social network, evaluating the three types of interactions.

1. Introduction

Recent changes in technology are radically changing the communication patterns, and how information reaches the vast majority of the population. The number of users engaged to online social networks, social media, or blogs, is rapidly growing all around the globe. Nowadays, Twitter is one of the most popular social media platforms and its main feature consists in allowing people to post and exchange text messages limited by 140 characters. This platform is specially suitable to conduct computational social science analysis [1], as it represents a wide variety of communications, going from personal to those coming from traditional mass media [2].

On Twitter all messages may be identified using keywords called hashtags [3]. This mechanism generates the trending topics, and people use them to discuss and exchange ideas without the necessity of having any explicit relation. Regarding research, the analysis of hashtag usage has helped to predict social relations [4] or collective attention [5].

In addition to hashtags, Twitter features several interaction mechanisms to facilitate the communication among users. These mechanisms establish different layers through which users can communicate and exchange information. Hence, Twitter can be seen as a multiplex or multilayer social network composed by the follower, mention and retweet layers. Multiplex networks [6–8] can help our understanding of a myriad of complex systems, ranging from social networks to biological systems, as most of them do not operate in isolation but through multiple interconnected layers. Thus, the main advantage of this

* Corresponding author.
E-mail address: fj.borondo@upm.es (J. Borondo).

http://dx.doi.org/10.1016/j.chaos.2014.12.023
0960-0779/

new formalism is that it incorporates multiple channels of connectivity, what makes it specially suitable to describe systems where the properties and neighbors of each node vary across layers.

In this paper, we present Twitter as a multilayer social network defined by the follower, mention and retweet interaction channels. The first interaction mechanism, is the ability of people to follow and be followed by the rest of users. This mechanism is a passive mechanism that allows users to receive the messages written by their followees at real time. By the same token they automatically deliver their posted messages to their followers. Thus, this mechanism establishes the followers layer, where users are connected among each other, according to who follows who. The links at this layer establish the substratum through which messages are delivered. Previous research have shown the complex properties in this network [9]. For example, it presents a scale-free [10] degree distribution, the small world effect [11,12], and a modular structure [13] with users clustered around leaders. Although having a large number of followers increases the visibility of the tweets posted by users, it not necessarily makes them influential [14]. Twitter also allows users to retransmit or retweet messages posted by someone else. The retweet mechanism allows individual messages to propagate and travel throughout the social network, and also serves as a way for people to endorse their point of view over specific subjects [15]. Finally, the third available interaction mechanism is the mention. By mentioning someone's username in the message text, people are able to send directed messages to other users. Whenever a user is mentioned on a tweet, he gets notified about it, as it appears in his private in-box, significantly increasing his chances of reading it. This mechanism is used either to establish conversations between users, through the exchange of messages; or to refer somebody in the messages text [16]. Overall, a high number of followers implies more visibility for the messages. However, it does not make a user influential in the active layers, as this depends on the value of the tweets content (retweets), or the name value of the user (mentions) [14].

In this paper we begin by characterizing the main properties of the three different Twitter layers: follower, mention, and retweet. For this purpose, we review our main findings related to two political conversations: the 2010 Venezuelan protest [17,18] and the Spanish general elections of 2011 [19,20]. We found that the structure of the follower layer conditions the retweet layer, as having a low number of followers represents a constrain to effectively propagate information on the retweet level. Next, we analyzed the rich-club ordering of the collapsed multiplex network. Moreover, we explored the influence gained by two differentiated type of accounts, *traditional media* and *politicians*, in the three available layers. Both types of accounts are supposed to have a high value name and should produce tweets with high value content. We found both of them to have a high visibility, as they were the top followed accounts. However, their influence on the active layers significantly differed. While politicians captured most of the collective attention, by having the highest in-degree in the mention layer; media accounts were the top influential on the retweet layer, as their accounts were the top retweeted. Finally, we analyzed how users clustered around these influential accounts in the three layers and show how the large and dense follower communities brake down into smaller and more segregated communities in the mention and retweet layers.

2. Dataset

The first dataset, regards the Venezuelan protest of 2010. This event took place exclusively on Twitter on December 16th, 2010. Two days before the protest, the convener asked his followers to post messages identified with the hashtag #*SOSInternetVE*. They massively responded and the conversation rapidly propagated becoming trending topic. This dataset was downloaded using the Twitter Search API version 1.0 to search for public access messages. This API provides data from a temporal index of recent tweets, posted within a lapse of a week from the time the query is made. The limitations of this API are not specified in terms of relative volume of tweets, nor a fixed number of queries. The dataset was built querying for messages with the hashtag #*SOSInternetVE*. We collected up to 421,602 messages, identified with the protest hashtag, which were posted by 77,706 users, between the 14th and 19th of December 2010 (two days before and after the protest). More details about the dataset can be found in our publications [17,18]. From now onwards we will refer to this conversation and dataset as *SOSInternetVE*.

The second considered dataset, relates to the 2011 Spanish general elections. To build this dataset we downloaded all the tweets using the Twitter API interface and searching for the specific keyword 20*N* in a three week period including the official electoral campaign and voting day. We chose this tag for being an ideologically neutral identifier, used all around Spain and by all the political parties when referring to this electoral process. More information about the dataset can be found in [19]. In what follows, we will refer to this dataset as 20*N*.

3. Twitter as a multilayer social network

By distinguishing among the different interaction mechanisms available on Twitter, we can define this platform as multilayer online social network of three layers: follower, mention, and retweet. We have illustrated this definition on Fig. 1. In it, the bottom layer (green) represents the follower layer, where links represent who follows who. The middle layer (pink), represents the mention layer. On it, links indicate who mentioned who on his tweets, and the weight of links represent the number of times it occurred. Finally, the top layer (blue) represents the retweet layer. At this level links indicate who retweeted whom, and the weight quantifies the number of times the retweet occurred. As the figure shows, not all users have to be present on all layers, although all of them will be present in the followers layer, by the sole fact of participating on the conversation. Similarly, links are not necessarily repeated on more than one layer, although mentions and retweets tend to occur through the followers layer. On

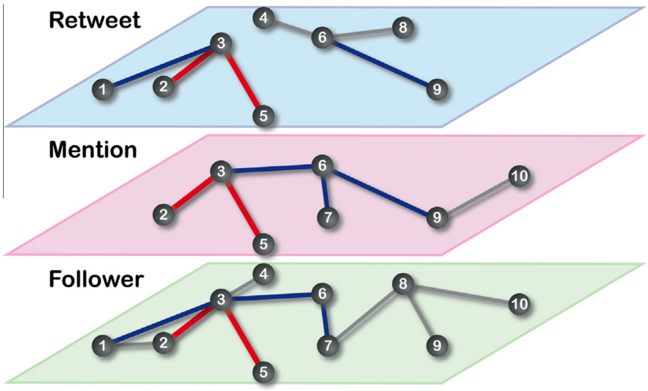

Fig. 1. Schematic representation of Twitter as a multilayer social network. The follower, mention and retweet layers have been represented at different levels. Links occurring on a single layer are colored in gray, while those occurring in two of them are colored in blue, and finally those present in all layers are colored in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the figure, gray links, indicate that the link took place on a sole layer, while those in blue occurred in two layers, and red ones were repeated in all the layers.

3.1. Follower layer

The first and most basic Twitter interaction layer is the Follower layer. We identify this layer as a passive layer that represents the social substratum through which most of the information flows. On this layer, a new edge is created whenever a user, A, decides to follow another user, B, (direction $A \rightarrow B$). Thus, the edges on this layer represent who follows who. In other words, the opposite direction of edges indicate who received whose messages and therefore, the direction in which information travels. As a consequence, the follow interaction is a nonreciprocal relation, and the network resulting from this layer is asymmetric with directed and non weighted links. At the end of the conversation the *SOSInternetVE* followers layer was compound by 77,706 nodes and 5,761,331 edges, while the 20*N* one was compound by 110,717 nodes and 6,031,076 edges.

The in-degree of a given node accounts for how many people follow him, i.e. how many people receive his messages. On the other hand, the out-degree measures the number of users that a certain user follows, which indicates from how many users he receives messages. Both the in and out degree distributions are presented on panels A, B, D and E of Fig. 2. For the Venezuelan conversation, the in and out degrees follow a heterogeneous distribution and can be fitted to power law distributions as noted in [17]. However, for the 20*N* conversation the out-degree distribution presents a less heterogeneous behavior, reaching lower maximum values. In terms of the in degree, the distributions indicate that over 50% of the users are followed by less than 15 users, while just around 1% of the users have over 1,000 followers. This fact, shows the existence of a minority of ultra connected users followed by a vast majority. A large number of followers, enhances the visibility of the messages posted by the user. However, it does not necessarily makes the user more influential, in terms of retweets or mentions gained. The presence of hubs with an extremely high in-degree or out-degree, together with the density of the network drives this layer to an average path length between 2 and 3 for the two conversations. This value indicates a small world [11] behavior or even an ultra small world [21]. This phenomenon was first reported by Stanley Milgram when he detected that on average two randomly chosen people could be linked through 6 hops (intermediary persons). Previous studies performed on the Twitter global follower graph state that the mean distance between users is around 4 [22]. However, our results are even lower than the previously reported values. This is because we are studying specific conversations, instead of the global Twitter Network. In this sense our samples correspond to an specific community of the global Twitter network, and therefore, the users engaged on it are closer among themselves than to the remaining users.

3.2. Mention layer

The second considered Twitter layer is the mention one. By mentioning someone's username in the message text, people are able to send directed messages to the mentioned user's inbox. This mechanism is often used to establish conversations between users, or just to refer somebody in the messages text [16]. Hence, in this layer a new edge appears when a new message posted by user, A, contains a mention to another user, B, (with direction $A \rightarrow B$). B is notified about the new tweet, what significantly increases his chances of reading it. In this layer the weight of links indicate the number of mentions going on among users.

Next, we computed the in and out strength cumulative distributions for both conversations, and plotted the results in panels A, B, D and E of Fig. 2 (green points). The in-strength indicates the number of mentions received by user. This measure indicates the fraction of the total collective attention that users gathered along the conversation. The in-strength distributions follow a power-law behavior for both datasets. As in the previous layer, the

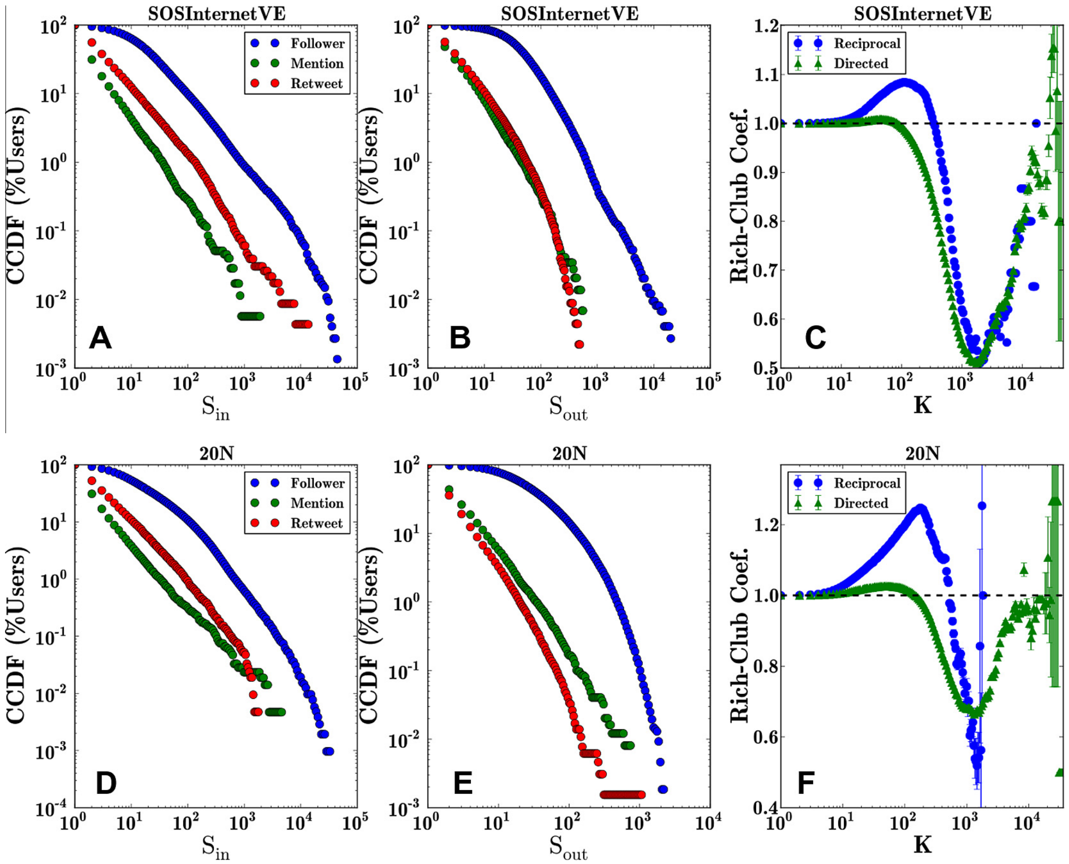

Fig. 2. (A) Complementary cumulative distribution of the in-strength at the follower (blue), mention (green) and retweet (red) layers for the *SOSInternetVE* dataset. (B) Complementary cumulative distribution of the out-strength at the follower (blue), mention (green) and retweet (red) layers for the *SOSInternetVE* dataset. (C) Rich-club coefficient for the directed (green) and reciprocal (blue) multiplex networks of the *SOSInternetVE* dataset. (D) Complementary cumulative distribution of the in-strength at the follower (blue), mention (green) and retweet (red) layers for the 20*N* dataset. (E) Complementary cumulative distribution of the out-strength at the follower (blue), mention (green) and retweet (red) layers for the 20*N* dataset. (F) Rich-club coefficient for the directed (green) and reciprocal (blue) multiplex networks of the 20*N* dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

collective attention was distributed in a highly heterogeneous manner. For example, for the 20*N* conversation, just 1.04% of the users captured 50% of the total mentions. On the other hand, the out-strength distribution indicates the amount of mentions posted by user. These distributions are not as heterogeneous as the in-strength ones, reaching lower maximum values. This is because they are limited by individual activity, while the in-strength results from a collective behavior. To further explore how users interacted with each other, we calculated the assortativity by degree coefficient (r) [23]. Since links are directed, we calculated this measure by splitting it into combinations of in and out degree pairs [24]. We found the layer to be slightly dissasortative. For example, for the 20*N* conversation the out-in pair took a value of $r_{out-in} = -0.141$. These results show the asymmetric shape of the layer, where the hubs that concentrate much of the incoming links are often mentioned by regular users who not mention frequently. This behavior is typical of online interactions [25], and is opposite to offline behavior. Offline social networks tend to be assortative, as popular people relate among them, and are unreachable for the mass. The explanation for this divergence of behavior between Twitter and the offline world, is that Twitter interactions largely differ from offline ones, in the sense that regular people are capable of reaching popular accounts, either by following, mentioning or retweeting them. Finally, we identified the top mentioned accounts finding that in their vast majority belonged to politicians or political party official accounts.

3.3. Retweet layer

The last considered layer is the retweet layer. The structure and heterogeneity of the follower layer has a big impact on retweets, as it raises a high level of disparity in the reception of the messages, and consequently in the information spreading process. This layer is considerably smaller and sparser than the followers one. This fact, evidences that users are much more selective when actively spreading information, than when just receiving or reading it [26]. To further understand how users retweet, we analyzed the emergent retweet network from the studied conversations. At this layer edges are created whenever a user retransmits a message originally posted by someone else. Hence, edges are directed and their weight indicates the number of times users retweeted each other, plus the number of subsequent propagators that retweeted the same message. Most of the flux at this layer occurs through links of the followers graph. This phenomenon is illustrated on Fig. 3A, where we have visualized a subset of the retweet layer (green edges), superimposed on the followers network (gray edges). In it, nodes who posted an original message are colored in red, while those who propagate it are colored in yellow. However, a fraction ($\sim 38\%$) of the total retweets were done by users not directly connected at the follower level to the original author of the tweet. The explanation for this behavior is that retweets tend to occur in cascades [18]. These cascades, emerge when a single message is transmitted by a user to his followers, subsequently allowing them and their own followers to do the same. We have sketched this phenomenon on Fig. 3B, that visualizes an example of a retweet cascade. On it, node 0 (colored in red) posts a new tweet. This tweet, travels through the follower layer to his followers (at 1 step distance from the source) enabling them to retweet the message. Nodes in white do not retweet it, while yellow ones do. Similarly, in the second time step, the tweet reaches the followers of the nodes who retransmitted it on the first instance (at 2 steps distance from the source), giving them the opportunity to retweet it or not. The process would continue until no more nodes retweet the considered message. Retweet cascades tend to be small, as more than half were formed by only two users besides the author and just a minority of them involve a large amount of users. The reason for this behavior is that information looses its attraction when farther from the authors social surroundings [27].

Next, we analyzed the number of retweets gained by user, R_{in}, which we label as the node in-strength of the retweet layer. The distributions can be found in panels A and D of Fig. 2 This measure quantifies the number of retweets gained by the user, regardless of whether he originally posted the tweet or he retweeted it. The retweet in-strength distribution follows a power law behavior for both conversations, as we already noted in [19,17]. For example for the *SOSInternetVE* conversation, only 25% of the overall users got retweeted at least once. Moreover, at this layer, the *SOSInternetVE* conversation was dominated by a minority of 0.4% of influential accounts, who concentrated around 50% of the retweets. The 20*N* conversation exhibited a similar behavior, as just 2.24% of the users gained over 50% of the total retweets. These influential accounts (for both conversations) predominantly belonged to traditional media.

These results show the difficulty of achieving a high number of retweets, and suggests that the majority of users would need to post an enormous amount of tweets to gain a significant number of retweets. Hence, we next address the following question: what is the relation between the influence users gain and the effort they must employ to do so? To answer this question, in [18], we proposed a measure to rank users according to their efficiency to propagate information. Accordingly, we defined *user efficiency*, η, as the ratio between the retweets gained by a user and the activity he employed for it. Thus, η can be expressed in the following way:

$$\eta = \frac{R_{in}}{A} \tag{1}$$

where A represents the user activity-the total number of messages he posted. Hence, $\eta = 1$ establishes the threshold from inefficient to efficient (more retweets gained that activity employed). In average most of the users who get retweeted, gain as many retransmissions as messages posted. However, a minority of them, occupying a privileged position in the followers network, accomplish a very

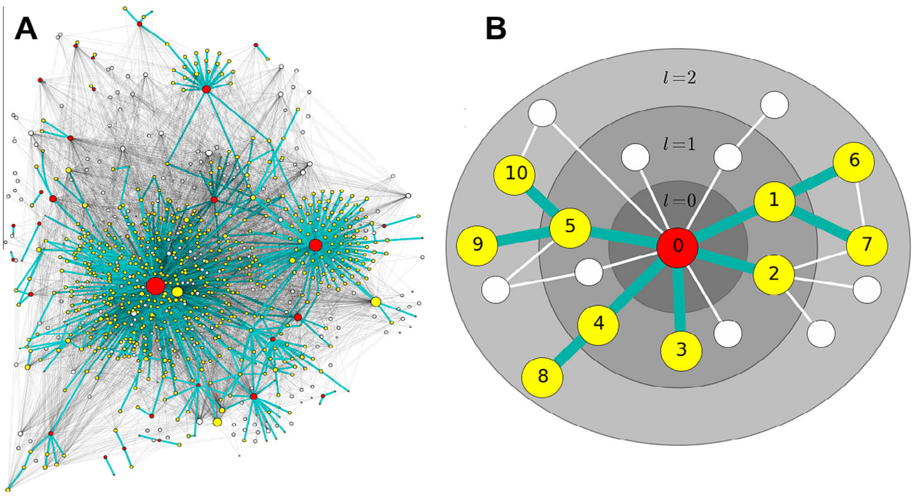

Fig. 3. (A) Visualization of a subset of the retweet layer superimposed on the follower layer for the *SOSInternetVE* dataset. Nodes in red posted an original message, and those in yellow retweeted a message. Links colored in green correspond to retweets, while those in gray to following relations. (B) Schema explaining how retweet occur in cascades. On it, we visually explain how the message posted by the red node travels to his followers allowing them to propagate it. In the next step, the followers of nodes who propagate the tweet receive it and are able to propagate it. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

high level of retransmission with little effort. This minority of extremely efficient users correspond to those with highest in-degree in the followers layer. This fact reflects that a lack of followers is a constrain to efficiently become influent at the retweet layer, while occupying a privileged position on the followers layer helps being highly retweeted. Hence, topocracy (the compensation for individuals is primarily determined by the position they occupy in a network) [28] seems to play a relevant role on Twitter. Moreover, in [18], we also introduced a model to control the effect that topology and different user activity strategies have on efficiency. We found that the appearance of a minority of highly efficient users results from the heterogeneity of the followers layer and independently of the individual user behavior.

4. Rich-club

Next, we explore whether the two studied Twitter conversations present a rich-club structure (i.e. highly connected nodes tend to connect among themselves) [29] or not. To this end, we measure the rich-club coefficient of the collapsed multiplex network-the aggregate of the three layers. A first definition of the rich-club phenomenon was introduced by Zhou and Mondragon [30] and can be expressed as:

$$\phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k} - 1)} \tag{2}$$

where $N_{>k}$ represents the number of nodes with degree higher than k and $E_{>k}$ denotes the number of edges among them. Hence, $\phi(k)$ measures the fraction between the number of actual edges and maximum number of edges that can exist among nodes with degree larger than k. This equation can be easily generalizable for directed networks, where we define the rich nodes as those with higher in-degree. Thus, Eq. 2 can be rewritten as:

$$\phi(k_{in}) = \frac{E_{>k_{in}}}{N_{>k_{in}}(N_{>k_{in}} - 1)} \tag{3}$$

However, hubs will be naturally more densely connected among themselves than nodes with lower degree. Thus, to properly interpret whether a network presents rich-club ordering we need to compare the rich-club coefficient with its randomized case. We randomize the networks to obtain uncorrelated networks with the same degree distribution of the original one. Hence, we can define the normalized rich-club coefficient ρ_{ran} as:

$$\rho_{ran}(k) = \frac{\phi(k)}{\phi_{ran}(k)} \tag{4}$$

where ϕ_{ran} is the rich club coefficient of the randomized network with the same degree distribution $P(k)$ of the original one. Note that for the directed case we preserved the in and out degree distributions. Values of ρ_{ran} larger than one indicate that the network presents a rich-club ordering, as the increase in the interconnectivity among large degree nodes is larger than what could be expected in the randomized case. In contrast, values below one evidence a lack of connectivity among hubs.

In order to analyze the rich-club ordering of Twitter we have first calculated ρ_{ran} for the total directed multiplex network of each conversation. Next, we have filtered these networks by only remaining reciprocal edges, and measured ρ_{ran} for the reciprocal cases. The results are presented in panels C (*SOSInternetVE*) and F (20*N*) of Fig. 2. The collapsed directed network does not present a rich-club structure. In fact, it presents a similar structure to the protein network [29], where hubs are not densely connected among themselves. This result, indicates that users with a high global in-degree are presiding large communities of regular users. This absence of rich-club ordering goes in agreement with the results presented in the previous sections and in [19]. In these sections, we reported the dissassortative nature of Twitter, where hubs with a large in-degree tend to be followed, metioned and retweeted by regular users. Despite, regular users can direct their attention to famous accounts on Twitter, these rich accounts do not interact with them. Hence, the reciprocal network does present a rich-club organization. The rich-club coefficient reaches its maximum around $k \sim 200$ and disappears for connectivities over 500. Hence, when considering reciprocal interactions the rich-club ordering emerges on Twitter and hubs preferentially interact among themselves in a similar way as elite scientists do in the scientific collaboration network [29].

5. Multiple leaders emerge at the different layers

In this section we explore whether the influence of two different elite collectives, such as politicians and mass media, is stable through layers, or if it varies across interactions. To this end, we first study the role played by politicians and mass media in each layer and how their influence varies across them. Following, we analyze the community structure of the layers, identifying the leaders of the top communities at each layer. Finally, we explore the existing relations among the communities of the different layers. To illustrate the results we will focus on the 20*N* dataset, although similar conclusions could be derived from the Venezuelan conversation following the same procedure.

We can begin to understand how the collective attention has been distributed in the three layers by identifying the top 50 influential users (those with higher in-degree) in each one. We chose the in-degree for being a measure of direct influence on Twitter. More particularly, we have studied the role played by politicians and traditional media on each layer and on the overlapping among layers. Thus, we have classified each of the top 50 influential accounts at each layer as either politician, media or blogger/others. We have visualized the results on Fig. 4 by representing the Ven diagram. On each region of the diagram we have indicated the percentage of accounts belonging to each collective: politicians, media, and bloggers. As the figure shows, the relevance of politicians and media varies according to the considered layer or layers. Overall, traditional media tend to be the most influential. However, when just considering the mention, or the overlap between the mention and follower layers, politicians are the most

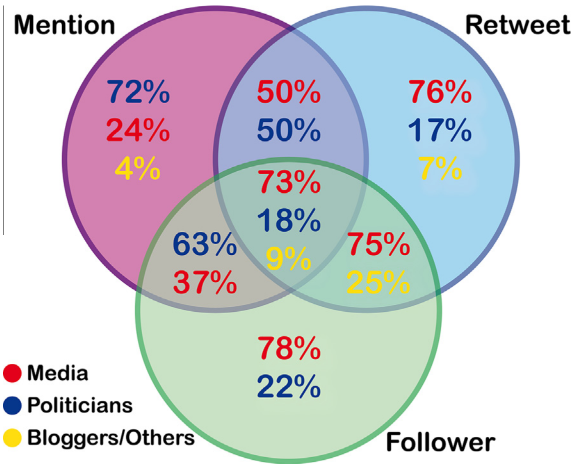

Fig. 4. Representation of the Venn Diagram for the follower, mention and retweet layers. The percentage of accounts in each region belonging to media (red), politicians (blue), and bloggers (yellow) have been indicated for the 20*N* dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

important. On the other hand, if just considering retweets, media emerge again as the top influential collective. While when considering the overlap between mention and retweet both elite seem to be equally popular. This highlights how conclusions can significantly vary from layer to layer, and therefore when just considering one layer these conclusions should limit to the considered layer, rather than general for the entire Twitter.

Next, to further understand the impact that media and politicians have on the structure of the different interaction layers, we analyze their community structure. For this, we have performed a community structure analysis (using the map equation algorithm [31]) of the follower layer and compare the results to those reported in [20] for the mention and retweet layers. Fig. 5 B shows the top nine communities of the follower layer, together with the main relationships among them. In this layer, communities are large and contain several influential accounts, related to a same collective-like Mass Media, Political Parties, Social Activism, or geographical region. For example, the largest community holds several important Spanish media (panel D). As it can be seen, various hubs stand out above the crowd. These hubs correspond to accounts of the main Spanish media, such as Europa Press, El Pais, or ABC. Another important community was formed around popular politicians and media from Catalonia (panel C). The main characteristic of this community is the use of the Catalan language. This community holds a majority of users ($\sim 66\%$) that preferentially tweeted in Catalan. In fact, the online Spanish political debate is segregated by language [20]. Other large communities clustered together users holding the similar political ideology. These communities were formed around a single political party accounts, and therefore exhibit a highly segregated partisan structure. Despite a small fraction of links across opposed ideology communities, users tend to interact with those holding a similar ideology. This phenomenon is illustrated in Panel E that visualizes the communities of the two dominant parties: Partido Popular (PP) and Partido Socialista Obrero Español (PSOE). In this panel PP has been colored in blue, while PSOE has been colored in red. As can be appreciated there is a high political polarization. This behavior is similar to that observed for the United States, on Twitter [32] and blogs[33].

Finally, we explore the relation among the community structure of the different layers. For this, we identified

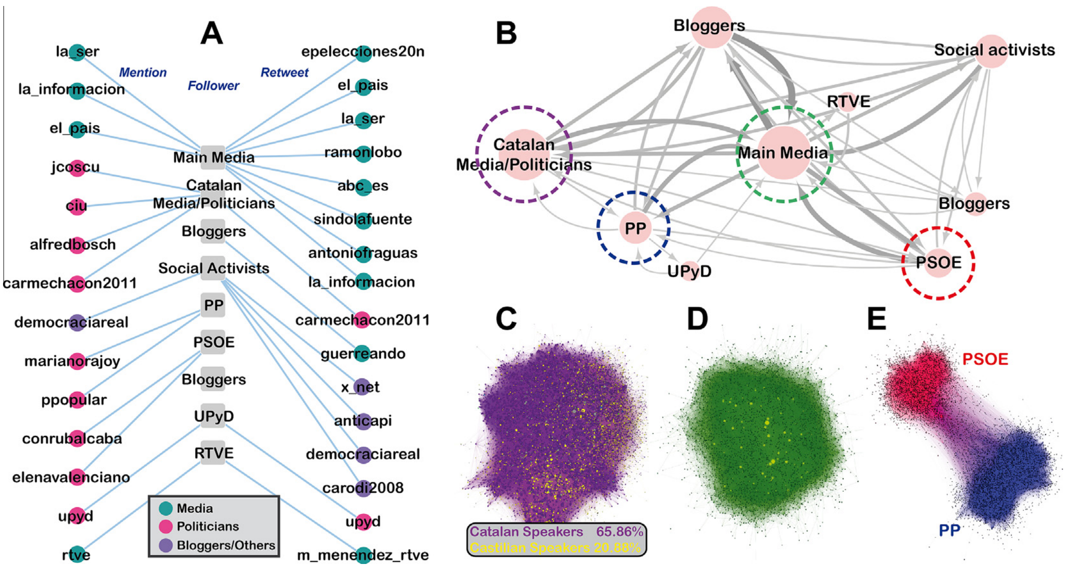

Fig. 5. For the 20*N* conversation: (A) Sketch, showing how the large communities of the follower layer split into several smaller communities in the mention and retweet layers. (B) Visualization of the network of communities in the follower layer. Nodes represent communities, and links account for relations among them. (C, D, E) Visualization of the inside structure of several communities at the follower level.

the top accounts of each community at the mention and retweet levels, and determined the community that they belonged in the follower level. We found that hubs embedded on large followers communities emerge as leaders of smaller and sparser ones in the mention and retweet layers. This showing that the large followers communities brake down into several smaller ones, formed around hubs belonging to a same follower community, at the active layers. Hence, users are more selective when extra effort is required to interact, holding less links and clustering in smaller and more selective groups. For example, while accounts belonging to different media may be classified on a same follower community, this phenomenon is not repeated in the retweet level. At this level media accounts belonging to a same community, also belong to a same media. This reflects that while users may passively follow different medias, they always relay on the same one to propagate information. This issue has been illustrated in Fig. 5 A, where we have sketched the main relationships between the followers communities and the mention and retweet communities. The links highlight how the top communities in the reweet (right side) and mention (left side) layers grow around high visible accounts belonging the top communities in the follower layer. In the visualization communities of the mention and retweet layers have been colored according to the collective leading the community (media in green, politicians in pink, and bloggers in purple). While most of the top communities in the retweet layer grow around media accounts, the top mention communities were formed around politicians. Hence, this showing again that politicians are the main characters in the mention layer, while traditional media accounts were the preferred source of information from which to propagate news at the retweet level.

All these showing that the communication patterns associated to each interaction mechanism are considerably different, what reflects the need to study Twitter as a multilayer network.

6. Discussion

In this paper we have defined Twitter as a multilayer social network, reviewed the main findings of our previous papers for each interaction layer, and highlighted how the influence of the elite varies across layers. For this matter, we have considered as cases of study the Venezuelan online protest of 2010 and the 2011 Spanish general elections.

The three differentiated Twitter interaction mechanisms, follower, mention, and retweet, define three layers through which individuals receive and diffuse information. Hence, the Twitter information diffusion process does not take place through a single channel, but three. In order to fully understand the process we have to simultaneously analyze all three channels. For example, the propagation of messages via retweets is strongly conditioned by the topology of the follower layer, as it establishes the substratum through which individuals receive information. Additionally, users establish conversations or refer to each other using the third available channel, the mention.

The collapsed directed multiplex network does not present a rich-club ordering, as politicians presided large communities of regular users in the mention layer; while media accounts were the sources from which people retweeted information. However, when considering reciprocal interactions the rich-club ordering emerges, as elite accounts preferentially interacted among themselves and largely ignored the crowd. The rich-club was mainly composed by politicians, media, and well-known bloggers. Hence, we identified the top 50 influential users at each layer, and classified them as media, politicians, or bloggers. Despite an slight overlapping among the top influentials at each layer, the relevance of the three different collectives significantly varied from one layer to another. The relevance of media and politicians at the follower level seems to be balanced. However, politicians clearly stand out in the mention layer, while media stand out in the retweet layer. A high degree in the mention layer is usually associated as a high value name, i.e. a famous and popular account, while the gain of retweets is associated to producing high value content tweets. Our results show that media were the sources of information, while politicians were the main characters of both conversations. Moreover, it suggests that politicians in general were not capable of producing high quality content tweets that got highly retweeted. All these resulted on users clustering around politicians in the mention layer, and around media accounts on the retweet layer. Hence, the leaders emerging at each layer vary significantly, and one can not claim neither politicians ruled the media or vice versa. It all depends on what kind of interactions we are considering and what effect we are trying to understand. Hence, we conclude that to fully understand Twitter we have to explore it as a multilayer social network, evaluating the three types of interactions.

Acknowledgments

This research was supported by the Ministry of Economy and Competitiveness-Spain under Grant No. MTM2012–39101.

References

[1] Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Others, Life in the network: the coming age of computational social science. Science 2009;323(5915):721 (New York, NY).

[2] Wu S, Hofman JM, Mason WA, Watts DJ. Who says what to whom on twitter. In: Proceedings of the 20th International Conference on World Wide Web, WWW '11. New York, NY, USA: ACM; 2011. p. 705–14. http://dx.doi.org/10.1145/1963405.1963504. <http://doi.acm.org/10.1145/1963405.1963504>.

[3] Pöschko J, Exploring Twitter Hashtags, CoRR abs/1111.6. <http://dblp.uni-trier.de/db/journals/corr/corr1111.html#abs-1111-6553>

[4] Romero DM, Tan C, Ugander J. Social-Topical Affiliations: The Interplay between Structure and Popularity, CoRR abs/1112.1. <http://dblp.uni-trier.de/db/journals/corr/corr1112.html#abs-1112-1115>

[5] Lehmann J, Gonçalves B, Ramasco JJ, Cattuto C. Dynamical Classes of Collective Attention in Twitter, CoRR abs/1111.1. <http://dblp.uni-trier.de/db/journals/corr/corr1111.html#abs-1111-1896>.

[6] Boccaletti S, Bianconi G, Criado R, del Genio C, Gómez-Gardeñes J, Romance M, et al. The structure and dynamics of multilayer networks. Phys Rep 2014;544(1):1–122. http://dx.doi.org/10.1016/j.physrep.2014.07.001. <http://www.sciencedirect.com/science/article/pii/S0370157314002105>.

[7] De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter M, et al. Mathematical formulation of multilayer networks. Phys. Rev. X 2013;3:041022. http://dx.doi.org/10.1103/PhysRevX.3.041022. http://link.aps.org/doi/10.1103/PhysRevX.3.041022.

[8] Cardillo A, Gómez-Gardeñes J, Zanin M, Romance M, Papo D, del Pozo F, et al. Emergence of network features from multiplexity, Scientific reports 3.

[9] Kwak H, Lee C, Park H, Moon S. What is Twitter, a Social Network or a News Media? Categories and Subject Descriptors, Most 112(2), 2010, pp. 591–600. arXiv:0809.1869v1, doi:10.1145/1772690.1772751. http://portal.acm.org/citation.cfm?doid=1772690.1772751

[10] Barabási AL, Albert R. Emergence of scaling in random networks. Science 1999;286(5439):509–12.

[11] Milgram S. The small world problem. Psychol Today 1967;2(1):60–7.

[12] Watts D, Strogatz S. Collective dynamics of small-world networks. Nature 1998;393:440–2.

[13] Newman MEJ. Modularity and community structure in networks. Proc Nat Acad Sci U S A 2006;103:8577–82. <http://arxiv.org/abs/physics/0602124>.

[14] Cha M, Haddadi H, Benevenuto F, Gummadi PK. Measuring user influence in twitter: the million follower fallacy. ICWSM 2010;10:10–7.

[15] Boyd D, Golder S, Lotan G. Tweet, Tweet, Retweet: conversational aspects of retweeting on twitter., In: HICSS, IEEE Computer Society, 2010, p. 1–10. <http://dblp.uni-trier.de/db/conf/hicss/hicss2010.html#BoydGL10>

[16] H. Courtenay, S.C. Herring, Beyond Microblogging: Conversation and Collaboration via Twitter, Tech. rep.

[17] Morales AJ, Losada JC, Benito RM. Users structure and behavior on an online social network during a political protest. Phys A Stat Mech Appl 2012;391(21):5244–53. http://dx.doi.org/10.1016/j.physa.2012.05.015. <http://www.sciencedirect.com/science/article/pii/S037843711200386X>.

[18] Morales AJ, Borondo J, Losada JC, Benito RM. Efficiency of human activity on information spreading on twitter. Soc Netw 2014;39:1–11. http://dx.doi.org/10.1016/j.socnet.2014.03.007. <http://www.sciencedirect.com/science/article/pii/S0378873314000215>.

[19] Borondo J, Morales AJ, Losada JC, Benito RM. Characterizing and modeling an electoral campaign in the context of Twitter: 2011 Spanish Presidential election as a case study. Chaos 2012;22(2):023138. http://dx.doi.org/10.1063/1.4729139. <http://www.ncbi.nlm.nih.gov/pubmed/22757545>.

[20] Borondo J, Morales AJ, Benito RM, Losada JC. Mapping the online communication patterns of political conversations. Phys A Stat Mechs Appl 2014;414:403–13. http://dx.doi.org/10.1016/j.physa.2014.06.089. http://www.sciencedirect.com/science/article/pii/S0378437114006529.

[21] Cohen R, Havlin S. Scale-free networks are ultrasmall. Phys Rev Lett 2003;90:058701. http://dx.doi.org/10.1103/PhysRevLett.90.058701. <http://link.aps.org/doi/10.1103/PhysRevLett.90.058701>.

[22] Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging usage and communities. In: Proceedings of the Ninth WebKDD and First SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis. ACM; 2007. p. 56–65.

[23] Newman MEJ. Mixing patterns in networks. Phy Rev E 2003;67(2):26126. http://dx.doi.org/10.1103/PhysRevE.67.026126.

[24] Foster JG, Foster DV, Grassberger P, Paczuski M. Edge direction and the structure of networks. Proc Nat Acad Sci 2010;107(24):10815–20.

[25] Hu H-B, Wang X-F. Disassortative mixing in online social networks. EPL (Europhys Lett) 2009;86(1):18003. <http://stacks.iop.org/0295-5075/86/i=1/a=18003>.

[26] Huberman B, Romero D, Wu F. Social networks that matter: Twitter under the microscope, First Monday 14 (1). <http://firstmonday.org/ojs/index.php/fm/article/view/2317>

[27] Wu F, Huberman BA, Adamic LA, Tyler JR. Information flow in social groups. Phys A Stat Mech Appl 2004;337(1):327–35.

[28] Borondo J, Borondo F, Rodriguez-Sickert C, Hidalgo CA. To each according to its degree: the meritocracy and topocracy of embedded markets. Sci Rep 2014;4(3784):1–4. http://dx.doi.org/10.1038/srep03784.

[29] Colizza V, Flammini A, Serrano MA, Vespignani A. Detecting rich-club ordering in complex networks. Nat Phys 2006;2(2):110–5.

[30] Zhou S, Mondragón RJ. The rich-club phenomenon in the internet topology. Commun Lett IEEE 2004;8(3):180–2.

[31] Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. Proc Nat Acad of Sci 2008;105(4):1118–23. http://dx.doi.org/10.1073/pnas.0706851105. <http://www.pnas.org/content/105/4/1118.abstract>.

[32] Conover MD, Ratkiewicz J, Francisco M, Gonc B, Flammini A, Menczer F. Political Polarization on Twitter. Networks 2011;133(26):89–96. http://dx.doi.org/10.1021/ja202932e. <http://truthy.indiana.edu/sitemedia/pdfs/conovericwsm2011polarization.pdf>.

[33] Adamic LA, Glance N. The political blogosphere and the 2004 U.S. election: divided they blog, In: Proceedings of the Third International Workshop on Link Discovery, LinkKDD '05, New York, NY, USA, 2005, p. 36–43.