

Research Article

Recurrent Patterns of User Behavior in Different Electoral Campaigns: A Twitter Analysis of the Spanish General Elections of 2015 and 2016

S. Martin-Gutierrez , J. C. Losada , and R. M. Benito 

Grupo de Sistemas Complejos, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Av. Puerta de Hierro, 2, 28040 Madrid, Spain

Correspondence should be addressed to J. C. Losada; juancarlos.losada@upm.es

Received 6 June 2018; Revised 16 September 2018; Accepted 29 October 2018; Published 5 December 2018

Academic Editor: Nam-Phong Nguyen

Copyright © 2018 S. Martin-Gutierrez et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We have retrieved and analyzed several millions of Twitter messages corresponding to the Spanish general elections held on the 20th of December 2015 and repeated on the 26th of June 2016. The availability of data from two electoral campaigns that are very close in time allows us to compare collective behaviors of two analogous social systems with a similar context. By computing and analyzing the time series of daily activity, we have found a significant linear correlation between both elections. Additionally, we have revealed that the daily number of tweets, retweets, and mentions follow a power law with respect to the number of unique users that take part in the conversation. Furthermore, we have verified that the topologies of the networks of mentions and retweets do not change from one election to the other, indicating that their underlying dynamics are robust in the face of a change in social context. Hence, in the light of our results, there are several recurrent collective behavioral patterns that exhibit similar and consistent properties in different electoral campaigns.

1. Introduction

Nowadays, social networking sites (SNS) are a well-established communication medium. They are used by a huge user base to share experiences, discuss opinions, read the news, etc. Twitter is one of the most dynamic SNS with respect to the interactions among users and one of the most powerful with respect to the potential information that can be extracted for research purposes. Moreover, in the last quarter of 2015 this social network had 305 million monthly active users, and the last data show that the number of monthly active users has risen to 335 million in the second quarter of 2018 [1].

These facts have stimulated the development of research projects from a wide variety of fields, from sociology to network science, that have provided new perspectives to the study of user behavior. These studies unravel the emergent patterns in our collective behavior and show how they can be used to gain insight about relevant social topics like economy [2], marketing [3], politics [4, 5], mobility [6], or polarization [7, 8].

In this work, we are interested in electoral processes. In these kinds of contexts, people use social media as a communication medium to exchange opinions. However, there are also some users whose purpose is to influence the conversation in a way that may affect the voting choices of the people. Although these are online actions, they have an impact on the offline world.

The concise character of a tweet turns it into a powerful tool to share breaking news and take part in dynamic debates [9]. That is why it is one of the most used media by journalists, politicians, and politics enthusiasts to share their views and news. Nowadays, even traditional media often use Twitter as a news source. Furthermore, politicians consider Twitter one of the most consequential social media for their job [10], although other online networking sites like Facebook also hold a high relevance.

In the Obama campaign of 2008, the efficacy of Twitter as a communication medium in a political context was established for the first time [11, 12]. From then on, it has gained more and more relevance in political campaigns. This

trend has also inspired an ever growing number of academic works that use Twitter data to analyze different aspects of the political campaigns.

Some lines of research are centered around predicting the outcome of elections following diverse techniques [4, 5, 13]. Other works investigate the interaction among parties to assess the cohesion of the parties [14] or the level of debate [5]. There are also studies that analyze the content of the tweets in order to detect their sentiment [15–17]. Two extensive reviews that cover a variety of subjects within the area can be found in [18, 19].

In a previous paper [5] the Spanish general elections of 2011 were studied using Twitter data. That work was focused on the prediction of the outcome of the election using a metric that measures political sentiment. The authors also explore the interactions among users by analyzing the structural and dynamical patterns of the complex networks emergent from the mention and retweet networks. From the study of the communication dynamics among politicians, they found a lack of debate, and a network growth model was proposed to reproduce these interactions.

Although we compare some of our results to those presented in that work, in this study we do not aim to predict the outcome of the elections or simply study the relations among users and politicians. We take advantage of the availability of Twitter data gathered during two electoral campaigns (the Spanish general elections of 2015 and the repetition of the elections in 2016) that are very close in time to compare the collective user behavior manifested in two analogous social systems with a similar context. The individual activity and the political actors may be different; in fact, in the second election two parties formed a coalition, altering the political landscape. Our objective is to study and characterize emergent behaviors that are recurrently manifested in political contexts.

To this end, we have computed and compared time series of daily activity for both elections, finding that the temporal series of activity for both electoral periods are significantly correlated. Furthermore, our results suggest that the number of tweets, retweets, and mentions follow a power law with respect to the number of unique users that take part in the conversation. In order to explore the evolution of the interactions among users, we have built networks of mentions and retweets and studied their temporal evolution. This has enabled us to verify that they show similar topological properties in both electoral periods. Besides, we have studied the mention and retweet subgraphs induced by political users and obtained results that imply a lack of communication among different parties and are in agreement with previous works [5]. Furthermore, we have found that an intensification of the interaction can be detected between parties after the formation of a coalition.

The paper is organised as follows. In the second section we describe the political context, some relevant aspects of the interaction mechanisms in Twitter, the characteristics of our dataset, and the methodology followed to build the networks of interactions. In the third section, we present and discuss our results with respect to the user activity and the evolution of the mention and retweet networks. We also

use a metric to discuss the influence of regular users and politicians. Furthermore, we analyze the degree of debate among politicians of different parties. Finally, we summarize the main conclusions.

2. Materials and Methods

2.1. Political Context. Spain has a bicameral parliamentary system, where the lower house is called Congress of Deputies and the upper house, the Senate. For elections to the Congress of Deputies, held every four years, each of the 50 provinces serves as an electoral district, with the number of deputies representing it determined by its population. Under a proportional representation electoral system governed by the d’Hondt formula, ballots are cast for a provincewide party list rather than for candidates representing individual constituencies.

About four-fifths of the members of the Senate are directly elected via a plurality system at the provincial level. Each province is entitled to four representatives; voters cast ballots for three candidates, and those with the most votes are elected. The remainder of the senators are appointed by the regional legislatures. Because representation is not based upon population, in the Senate smaller and more-rural provinces generally are overrepresented in relation to their overall population [21]. Hence, with the purpose of illustrating the Spanish political landscape, in this work we have chosen to report only the results for the Congress of Deputies, as the distribution of seats is supposed to be more representative.

On the 20th of December 2015, the Spanish general elections were held. The PP (Partido Popular, People’s Party) and the PSOE (Partido Socialista Obrero Español, Spanish Socialist Workers’ Party), which constituted the traditional two-party system had lost a lot of social support, while the emerging parties Podemos (We Can) and Cs (Ciudadanos, Citizens) were on the rise. This caused a transition from a two-party system to a multiparty system [22].

In spite of that, the PP, which was holding the government, was still leading the polls. The rest of the parties were behind but not too far away. In fact, the supports for the other three main parties fluctuated so much during the year before the election that it seemed impossible to predict, by looking at the polls, which would be the final ranking of votes [23].

The result was a fragmented parliament where no party held an absolute majority and large coalitions were needed to form a government. The votes and seats that each party obtained are displayed in Table 1. After several months of negotiations there was no agreement between any group of parties large enough to obtain a parliamentary majority that would allow the formation of a government. This led to the announcement of a new election on the 26th of June 2016.

Before this new election, one of the emerging parties, Podemos, formed a coalition with IU (Izquierda Unida, United Left). This alliance was called UP (Unidos Podemos, United We Can).

The 2016 election resulted in a parliament that was almost as fragmented as the one in 2015. The votes and seats obtained by each party are presented in Table 1. In October 2016, the

TABLE 1: Results of the Spanish general elections of the 20th of December 2015 and the 26th of June 2016. Podemos and IU are together in 2016 because they formed a coalition called UP.

	Votes		Seats	
	2015	2016	2015	2016
PP	7236965	7941236	123	137
PSOE	5545315	5443846	90	85
Podemos	5212711	5087538	69	71
IU	926783		2	
Cs	3514528	3141570	40	32
Others	2775011	2665069	26	25

PP finally won the appointment vote and formed a minority government (in June 2018, a motion of no-confidence was issued against the ruling party (PP) and PSOE formed a minority government).

2.2. Description of the Data. We have worked with Twitter messages retrieved with the Twitter Streaming API. This API allows downloading *tweets* matching a set of keywords. In order to avoid biased results, we have chosen the following neutral keywords to filter the messages:

- (i) Keywords for the 2015 election: *20D*, *20D2015*, *#EleccionesGenerales2015*
- (ii) Keywords for the 2016 election: *26J*, *26J2016*, *#EleccionesGenerales2016*, *#Elecciones26J*

We have downloaded tweets during a period of more than two months before and after each election. However, the core of our analysis has been focused on the 15 days of the official electoral campaign, the *reflection day* (day before the election), the election day, and the day after. During that period of 18 days we have retrieved 1793145 *tweets* for the 2015 election and 1755438 for the 2016 election.

Besides the keywords used to retrieve the data, the most used hashtags that appear in our dataset are the following:

- (i) Top hashtags of 2015: *#podemos*, *#psoe*, *#partidopopular*, *#ciudadanos*, *#l6elecciones*, *#7deldebatededecisivo*, *#votapsoe*, *#possible*, *#pp*, *#mivotocuenta*, *#españa*, *#españaenserio*, *#podemos20dic*, *#20dicpodemos*, *#podemosremontada*, *#votapp*, *#hevotado*, *#votapodememos20d*
- (ii) Top hashtags of 2016: *#afavor*, *#unidospodemos*, *#l6elecciones*, *#votapsoe*, *#Debate13j*, *#psoe*, *#cambioamejor*, *#votapp*, *#españa*, *#ciudadanos*, *#partidopopular*, *#elcanvipossible*, *#unsiporelcambio*, *#avotar*, *#lasonrisadeunpais*, *#brexit*, *#eleccionesgenerales*

We have also compiled lists of Twitter accounts associated with the four main parties (PP, PSOE, Podemos and Cs) and with IU. The latter is relevant because, in the 2016 election, as we explained in Section 2.1, this party formed a coalition (UP) with Podemos, a fact that is reflected on the data. These lists of accounts allow us to analyze the difference of behaviors of regular users and politicians.

In order to build the set of political accounts, we have looked into the Twitter lists (which are lists of accounts

elaborated by the users) defined by relevant official accounts associated with each party. We have downloaded those lists that include politicians, political institutions, or supporters of the party. The total number of retrieved political users that participated in the conversation was 5227 in 2015, with an average number of followers of 4044, and 5012 in 2016, with an average number of followers of 4662.

2.3. How Does Twitter Work? In this section we will briefly describe some of the characteristics of Twitter. Specifically, the means of communication among users, which will be the source used to build the networks of interactions.

In Twitter there are several mechanisms of interaction among users. The first one is to *follow* other users. A user will receive every *tweet* that her followers post. Besides, there are five other mechanisms of interaction within the social network: the direct message, the *mention*, the *retweet*, the *quote*, and the *reply*. We will focus on the *mentions* and the *retweets*, as they are the most widely used.

The *mention* is a public direct communication mechanism. It consists of including the name of a user in a *tweet* ([..]@username[. .]). This way, the mentioned user will receive the message even if she is not following the user that posted it. There exists a convention in Twitter according to which, whenever a given user is mentioned inside a tweet, her @username is used instead of her real name, independently of whether the intention is to establish a communication with the target user or not.

The *retweet* is a broadcast mechanism. Every retweeted message by a user is forwarded to her followers the same way as an original tweet. In a political context retweeting usually, but not always, implies the endorsement of the ideas contained in the message.

2.4. Building the Networks of Mentions and Retweets. Here we will explain the methodology that we have adopted to build the networks of interactions from Twitter data. The interactions that we have analyzed are the *mention* and the *retweet*.

We have built mention networks by considering each user participating in the conversation as a node. Two nodes j and i are joined with a directed link from j to i when user j mentions user i ; that is, j posts a *tweet* including @name_of_i. The result is a directed and weighted network where the weight of each link corresponds to the total number of tweets

where j mentions i gathered during a given interval of time. The in-degree, or k_{in}^i , of a user i in this network corresponds to the total number of users that have been mentioned i . k_{out}^i corresponds to the total number of users that she has mentioned.

The retweet network is built in a similar way as the mention network. The nodes are the users and user j is joined with a directed link to user i if j retweets a message *originally* posted by i . If a user j retweets an original tweet by i and k retweets the message broadcast by j , there are links from k to i and from j to i , but not from k to j . Figure 1 illustrates this mechanism.

Let us remark this idea: links in the retweet network join retweeters with original posters; there are no links between the middlemen that broadcast the original message. The reason for this choice of methodology is that the Twitter API used to download the data at that moment only provided information about the original poster of a tweet in the retweet metadata. The retweet network is then a directed and weighted network where the weights of the links from j to i are the number of messages of i retweeted by j during a given period of time. k_{in}^i of a user i in this network corresponds to the total number of users that have retweeted messages by i . k_{out}^i corresponds to the total number of users whose messages she has retweeted. The retweet mechanism usually results in networks that display large star subgraphs.

We have considered two different temporal scales. On one hand, we have built daily networks, counting the 24 hours of a day from 4 AM (UTC) (in Spain, the local time corresponds to UTC+1 for winter time and UTC+2 for summer time) of that day until 4 AM of the following day. This way we capture full *human* activity cycles. This approach is based on a work by Morales et al. [24], where human synchronization is studied at several scales, finding strong periodicities linked to day-night and social cycles as well as a developing global synchronization that adds a new perspective to the phenomenon of globalization. On the other hand, we have built aggregated networks for the 15 days of electoral campaign plus the next three days (before election, election, and after election).

3. Results

3.1. Temporal Evolution of User Activity. In order to characterize the global activity of the users (number of tweets posted during a given period of time), we have computed time series of total daily activity for the whole period considered. Additionally, we have analyzed the temporal evolution of the distribution of daily activity per user.

In the left panels of Figure 2 we have plotted the time series of daily activity for both electoral campaigns. We have aligned them such that the voting day coincides with day 0. As we can see in that figure, before the electoral campaign the time series of daily activity present low values, but they start to rise around the onset of both campaigns.

In line with the existing literature [5, 25, 26], the features of the time series of activity can be explained to a high extent by offline events. There is a huge peak of messages on the election day, which accounts for more than 30% of the total of

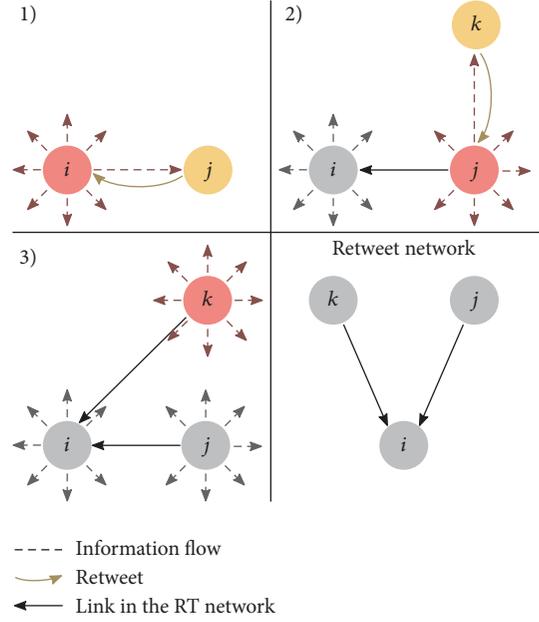


FIGURE 1: Scheme that shows the method used to build the networks of retweets. In panel 1), user i posts a message starting an information flow that reaches user j , who decides to retweet it. In panel 2), the message retweeted by j is read by k who in turn retweets it again. In panel 3) the three users are information transmitters and in the last panel the resulting retweet network is shown.

messages during the period of study (36% in 2015 election and 32% in 2016). There are also smaller peaks during the period of the electoral campaign that can be explained by the main electoral debates.

In Spain, preelectoral silence is mandatory during the period that spans from 12AM of the day before the election to the closing of voting polls. This period is called *reflection day* and a number of restrictions apply with respect to the behavior of political parties and media. Among others, electoral events can not be held and the spread of electoral propaganda is forbidden [27]. The time series experiences a decrement that day, which may be triggered by the fact that politicians cease to campaign in both online and offline media. Because of this lack of stimulus, the conversation among common users decays.

It is worth pointing out the similarity of the temporal evolution of the user activity in both elections. In Figure 2, we have plotted (in logarithmic scale) the time series of one period against the other such that the election days coincide. By performing a linear regression of these data, a significant linear correlation has been found. We obtain $r^2 = 0.99$ when we perform the linear regression including the day of the elections and $r^2 = 0.84$ if we exclude it. However, notice that the slopes present only a difference of 3.5%. This correlation can be partially explained by the similarity of the contexts and by the recurrent structure of the electoral campaign. It always starts 16 days before the election and lasts for 15 days. Moreover, the main debates were both held 13 days before the election. Due to the fact that the number of tweets in the 2016

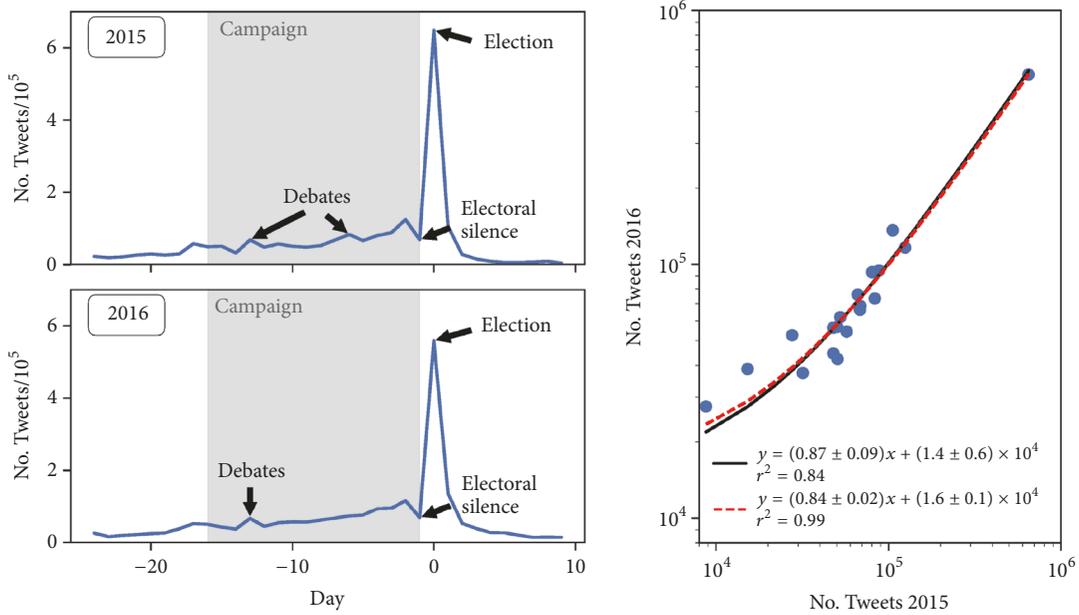


FIGURE 2: Left panels: time series of aggregated user activity per day for the 2015 (top) and 2016 (bottom) elections. The shadowed region corresponds to the days of electoral campaign. Right panel: linear regression between the activity time series of both elections. The lines correspond to the linear fits of the data including the day of the elections (red dashed line) and excluding it (black continuous line).

election was slightly lower than in 2015, the slope of the linear regression is lower than 1.

Another relevant property of the user activity is displayed in Figure 3, where we show that the total number of tweets, retweets, and mentions per day follows a power law with respect to the number of unique users each day. Accordingly, we have fitted our data to a power law with exponent α . The values of the exponent α and correlation coefficients of the fits are displayed in Table 2. To understand this behavior we have followed the work by Leskovec et al. [28], where it is shown that when real-world networks evolve through time, the number of links $E(t)$ scales with the number of nodes $N(t)$ as

$$E(t) \propto N(t)^\alpha \quad (1)$$

Notice that, whereas in 2015 the growth for the three quantities was slightly super-linear with respect to the number of users, in 2016 we observe an approximately linear behavior. Hence, in 2015 when more users join the conversation, the activity experiences a proportionally higher increment than in 2016.

In order to further explore the characteristics of the user behavior, we have also analyzed the temporal evolution of the distribution of the daily user activity shown in the left panels of Figure 4. The activity distributions show a heterogeneous character, result that is in concordance with the literature [29]. We have fitted them to the following discrete power law:

$$P(x) = \frac{x^{-\gamma}}{\zeta(\gamma, x_{min})}; \quad (2)$$

$$\zeta(\gamma, x_{min}) = \sum_{x=0}^{\infty} \frac{1}{(x + x_{min})^\gamma}$$

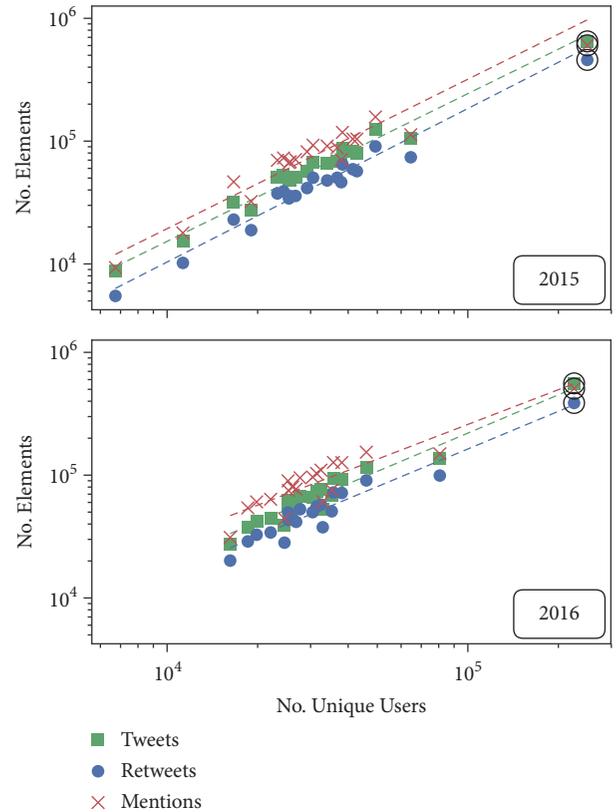


FIGURE 3: Power law relationships of the total number of tweets, retweets, and mentions per day as a function of the number of unique users that participated in the conversation each day for the 2015 campaign (top) and the 2016 campaign (bottom). Note that the data corresponding to the day of the elections (marked with a circle) were not included in the fits.

TABLE 2: Values of the exponents (α) and correlation coefficients (r^2) of the linear regressions of the log-log plots of tweets, retweets, and mentions with respect to the number of unique users per day shown in Figure 3 for both electoral campaigns.

	α		r^2	
	2015	2016	2015	2016
Tweets	1.20 ± 0.06	1.04 ± 0.11	0.96	0.85
Retweets	1.25 ± 0.08	1.02 ± 0.12	0.94	0.80
Mentions	1.22 ± 0.10	0.9 ± 0.2	0.89	0.66

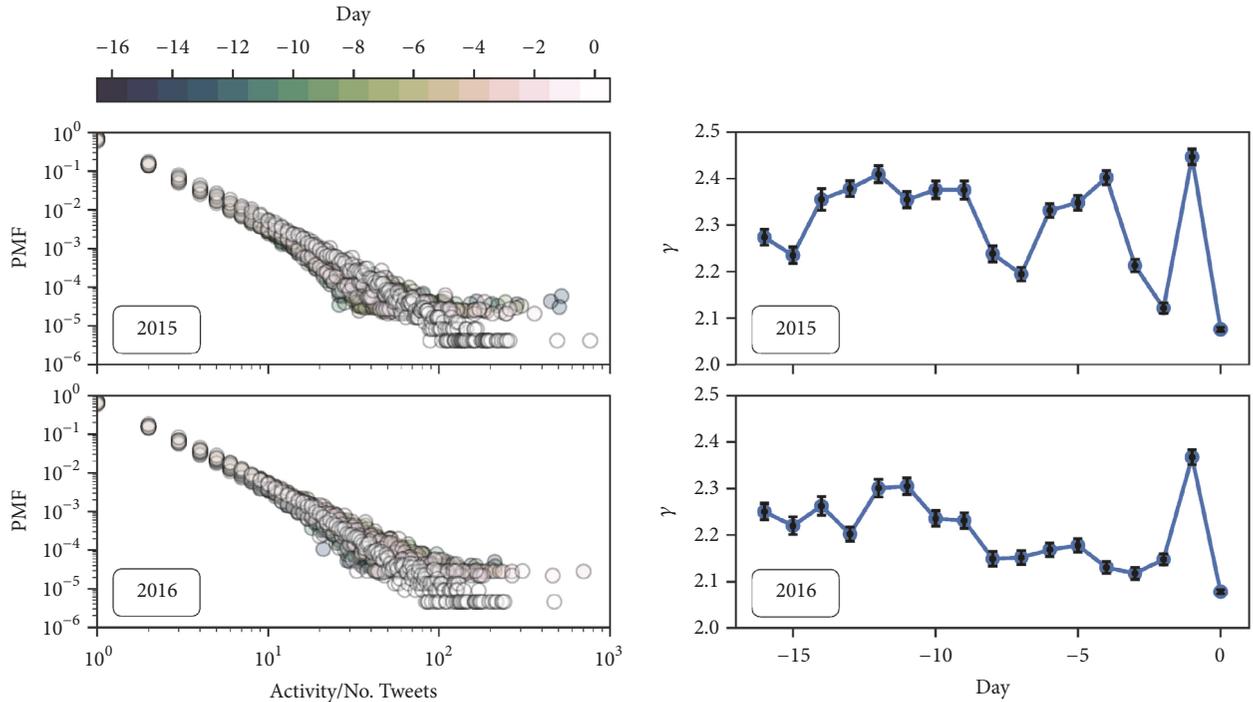


FIGURE 4: Temporal evolution of the distribution of activity per day in both electoral campaigns (2015 in the top panels and 2016 in the bottom panels). Left panels: probability mass functions (PMF) of the distribution of activity for each day in color code. Right panels: daily evolution of the γ exponent of the power law fit of the activity distributions. The error bars correspond to 2σ .

where in this case x is the daily activity and ζ is the Riemann zeta function.

In the right panels of Figure 4 we have plotted the temporal evolution of the exponent γ of the power law. It has been computed by numerically solving the following equation to obtain the maximum likelihood estimator [30] (MLE) of γ :

$$\frac{\zeta'(\gamma, x_{min})}{\zeta(\gamma, x_{min})} = -\frac{1}{n} \sum_{i=1}^n \ln(x_i) \quad (3)$$

where the prime denotes differentiation with respect to the first argument and n corresponds to the number of users. The uncertainty of γ has been computed as follows:

$$\sigma = \frac{1}{\sqrt{n [\zeta''(\gamma, x_{min}) / \zeta(\gamma, x_{min}) - (\zeta'(\gamma, x_{min}) / \zeta(\gamma, x_{min}))^2]}} \quad (4)$$

This technique is proven to be more precise than a minimum squares fit to the log-log plot, which usually yielding incorrect results when computing the parameters of power law distributions [30]. Every other exponent of power laws of discrete data has been computed in the same way. Since the minimum value of activity is 1 (and is also the most abundant in the data), we have fixed $x_{min} = 1$ when computing the MLE of γ .

We can see that the values of the γ exponent fluctuate within a small interval ($2.0 < \gamma < 2.5$) during the considered period for both elections. These values are perfectly compatible with those presented in [5], where the authors obtained a value of $\gamma = 2.275 \pm 0.002$.

The small fluctuations of the γ exponent tell us that the collective behavior does not change much from one day to the other. If we take into account the fact that the value of the exponent of the distribution controls how fast

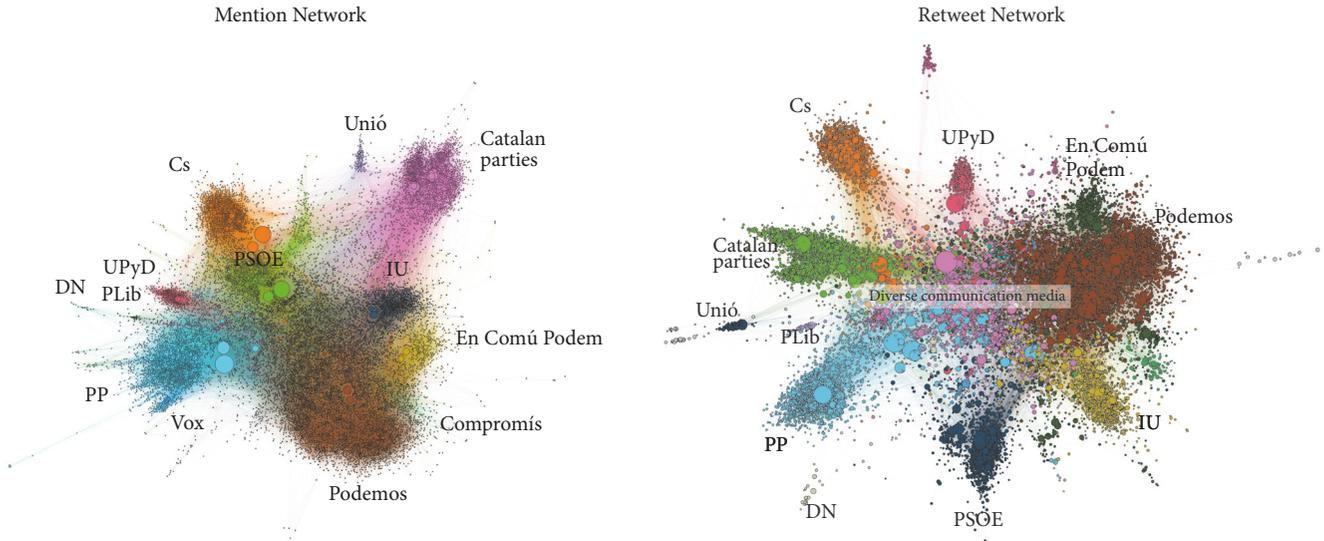


FIGURE 5: Left panel: strongly connected component of the aggregated mention network for the 2015 electoral campaign. Right panel: strongly connected component of the aggregated retweet network for the 2015 electoral campaign. Colors correspond to the communities computed with the Louvain algorithm [20]. We have indicated the most probable affiliations of the nodes of each community by visually inspecting which users correspond to the most central nodes. Every well-defined group seems to correspond to a political party. The size of the nodes is proportional to $\log(\text{PageRank})$.

it decays, we see that, when the exponent is smaller, the activity reaches higher values (and vice versa). Consequently, by looking at Figure 4, we can see that the day of the elections we obtain a low exponent due to the increase in activity. Additionally, during the electoral silence, consistently with the results shown in Figure 2, we appreciate a decrease in activity. Notice however that while in the case of Figure 4 the points corresponding to the electoral silence are the highest, they are not the lowest in the activity time series of Figure 2. This implies that although individually each user tends to post fewer messages that day, there are still a lot of users taking part in the conversation.

3.2. Temporal Evolution of Mention and Retweet Networks. We have analyzed the temporal evolution of the aggregated mention and retweet networks at two different temporal scales. On one hand, we have aggregated the networks for the whole campaign period (plus the next three days); on the other hand, we have performed an analysis of the temporal evolution of the networks by aggregating the data for each day separately and computing time series for different metrics.

In Figure 5 we show the strongly connected component of the aggregated networks of 2015. Colors correspond to the communities computed with the Louvain algorithm [20]. We have obtained a modularity of $Q = 0.72$ for the retweet network and $Q = 0.67$ for the mention network. Other community detection algorithms were applied obtaining analogous results [31]. We have indicated the most probable affiliations of the nodes of each community. In order to do that we have visually inspected which nodes (or users) are the most central in each community. The centralities of the nodes have been computed with the *PageRank* metric [32], which measures the influence of a node based on its neighborhood.

The sizes of the nodes in Figure 5 have been represented proportional to $\log(\text{PageRank})$.

Every well-defined group seems to correspond to a political party. Whereas in the representation of the mention network the different groups are well-defined and the most central nodes correspond to the leaders of the communities and are politicians or political parties, in the retweet network a mixing of nodes of different communities placed in the center of the representation can be appreciated. Most of these nodes present high centralities and belong to different communication media. This is in good agreement with the literature [33].

We have computed several statistical properties of these networks: the number of nodes (N) and links (E), the density ($\rho = E/(N(N-1))$), the average clustering coefficient (\bar{C}), and the in- and out-degree distributions. We have found that they are very heterogeneous and fitted them to a power law following the methodology described in Section 3.1. The in-degree (k_{in}) distribution is more heterogeneous than the out-degree (k_{out}) distribution; that is, the γ exponents follow the relation $\gamma_{out} > \gamma_{in}$. This is due to the fact that the k_{out} is associated with individual efforts whereas the k_{in} corresponds to collective actions. An individual is physically limited to posting a given number of tweets during a given period of time, whereas a large group of users may post many more messages in the same period of time. Because of this, high k_{in} values are more probable than high k_{out} values, leading to the aforementioned relationship between the γ exponents.

The properties of the aggregated networks are displayed in Table 3, where we can appreciate clear differences between the mention and the retweet networks. First of all, the former presents higher average and maximum degree. Whereas the higher γ_{in} of the mention network shows that its k_{in}

TABLE 3: General statistical properties of the retweet and mention networks aggregated for the whole period of study for both electoral campaigns. $k_{in(out)}$ corresponds to in- (out-) degree and $\gamma_{in(out)}$ to the exponent of the correspondent degree distribution. The confidence intervals in the power law exponents calculations correspond to 2σ .

Statistics	Mention		Retweet	
	2015	2016	2015	2016
Nodes	354079	319961	330071	296645
Links	1361495	1251113	928492	852216
Density	1.09E-05	1.22E-05	8.52E-06	9.68E-06
$\overline{k_{in}} = \overline{k_{out}}$	3.85	3.91	2.81	2.87
$(k_{in}^{min}, k_{out}^{min})$	(0, 0)	(0, 0)	(0, 0)	(0, 0)
$(k_{in}^{max}, k_{out}^{max})$	(17747, 824)	(16829, 645)	(9054, 459)	(8985, 534)
γ_{in}	1.606 ± 0.005	1.623 ± 0.005	1.587 ± 0.005	1.577 ± 0.006
γ_{out}	1.827 ± 0.003	1.825 ± 0.003	2.034 ± 0.004	2.022 ± 0.004
\overline{C}	0.19	0.19	0.06	0.06

distribution is slightly less heterogeneous than the one for the retweets, the lower value of γ_{out} implies that the k_{out} distribution is more heterogeneous for mentions than for retweets. The first phenomenon can be attributed to the construction methodology of the networks. Since all retweeters are linked to the original poster, high values of k_{in} are slightly more common in the retweet network. The second phenomenon arises from the fact that there can be more than one mention per tweet and in every retweet there is, at least, one mention to the original poster.

The higher average clustering (\overline{C}) for mention networks can be explained by the different uses of mentions and retweets. Mentions are a communication and allusion mechanism whereas the retweets are used mainly to broadcast messages, leading to more clustered networks on the first case and to more star-like subgraphs on the second one. Nevertheless, the construction mechanism of the networks also plays a relevant role, since the relationships between middlemen are not present in the retweet network.

The results show very little change from one election to the other, suggesting that the underlying dynamics of the networks of interactions are consistent and, to some extent, independent of the context.

With respect to the temporal evolution of the exponent of the in-degree distribution for mention and retweet networks, which is shown in Figure 6, it remains approximately constant for 47 days before the election (for simplicity, only 16 days before the election are shown in the figure). The average values during that period are $\gamma_{in}^{2015} = 1.68 \pm 0.08$ and $\gamma_{in}^{2016} = 1.67 \pm 0.07$ for the mention networks and $\gamma_{in}^{2015} = 1.65 \pm 0.08$ and $\gamma_{in}^{2016} = 1.63 \pm 0.07$ for retweet networks, with a confidence interval of 2σ . These results indicate a high consistency of the user behavior for both elections. The exponents are slightly higher to the one of the aggregated networks because in those networks almost everyone has more mentions, making the tail of the distribution longer and heavier.

In the case of the temporal evolution of the out-degree distribution exponent, displayed in Figure 6, it also remains approximately constant for 47 days before the election. The average values during that period are $\gamma_{out}^{2015} = 1.99 \pm 0.09$ and

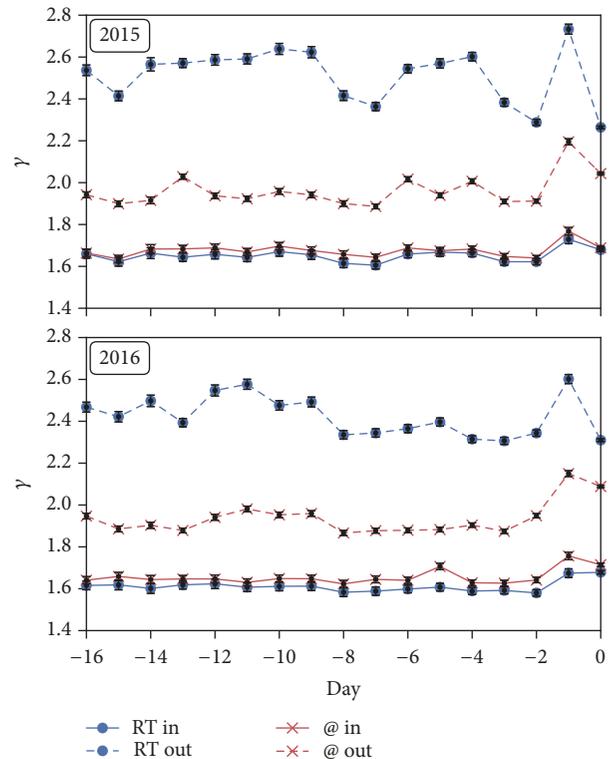


FIGURE 6: Daily evolution of the γ exponents of the power law of the in- and out-degree distributions for mentions (@) and retweets (RT) networks for the 2015 (top) and 2016 (bottom) electoral campaigns.

$\gamma_{out}^{2016} = 2.00 \pm 0.08$ for the mention networks and $\gamma_{out}^{2015} = 2.7 \pm 0.3$ and $\gamma_{out}^{2016} = 2.6 \pm 0.2$ for the retweet networks, with a confidence interval of 2σ . Analogously to the in-degrees, the similarity of the values of the exponents from one year to the other suggests a recurrent behavioral pattern. They are also similar to the values of the aggregated network, but higher, meaning that the heterogeneity is lower. The reason is again the construction mechanism of these networks. People are active during all the period under study, producing more links each day, thus slightly changing the distribution to a more heterogeneous one as more days are aggregated.

TABLE 4: Assortativities of the different aggregated networks.

Link direction	Mention		Retweet	
	2015	2016	2015	2016
out-in	-0.1096	-0.1095	-0.1234	-0.1169
in-in	-0.0174	-0.0162	-0.0313	-0.0287
in-out	-0.0006	-0.0014	0.0109	0.0074
out-out	0.0521	0.0488	0.0988	0.0892

TABLE 5: Z-Scores of the assortativities of the different aggregated networks with respect to 500 realizations of the directed configuration model.

Link direction	Mention		Retweet	
	2015	2016	2015	2016
out-in	-152	-137	-141	-133
in-in	-22	-19	-31	-31
in-out	2	2	12	9
out-out	100	95	107	113

We have explored the degree correlations of the networks and their evolution in several ways. First, we have computed the degree assortativities [34], presented in Table 4, and compared them with assortativities of randomized networks via a Z-Score in order to determine if the assortativities are a result of the degree distributions alone or of more complex correlations. The Z-Scores can be consulted in Table 5. The randomization has been carried out using 500 realizations of the directed configuration model [35] implemented in the NetworkX Python module [36].

As we can see in Table 4, the assortativities display a recurrent order among networks and years: $r_{out-in} < r_{in-in} < r_{in-out} (\approx 0) < r_{out-out}$. Consistently with the rest of the measures presented so far, they exhibit a high similarity among years. Moreover, when we compare the results of the degree correlations of the retweet network to those of the 2011 election [5], we see that they are very similar. This result suggests a recurrent behavior in the communication patterns of the users.

This order is maintained for the Z-Scores (see Table 5), meaning that the assortative relations are still more assortative than what we should expect for a random network and that the disassortative ones are more disassortative.

With respect to the temporal evolution of the daily degree assortativities that is displayed in Figure 7, we can see that they fluctuate around a more or less fixed value. However, these fluctuations are sometimes strong enough to alter the previously described order.

These fluctuations seem to be caused by disruptive events, both exogenous and endogenous. On day -6 of 2015 a debate was celebrated between the leader of the PP (then, the president) and the leader of the PSOE. A similar pattern, but weaker, can be observed in the time series of assortativity on day -13, which coincides with the other important debate of the campaign. On day -11 of 2016 we have found a viral tweet that contained a comical video about Spanish politics. The retweets received by that particular tweet amounted to 4.7%

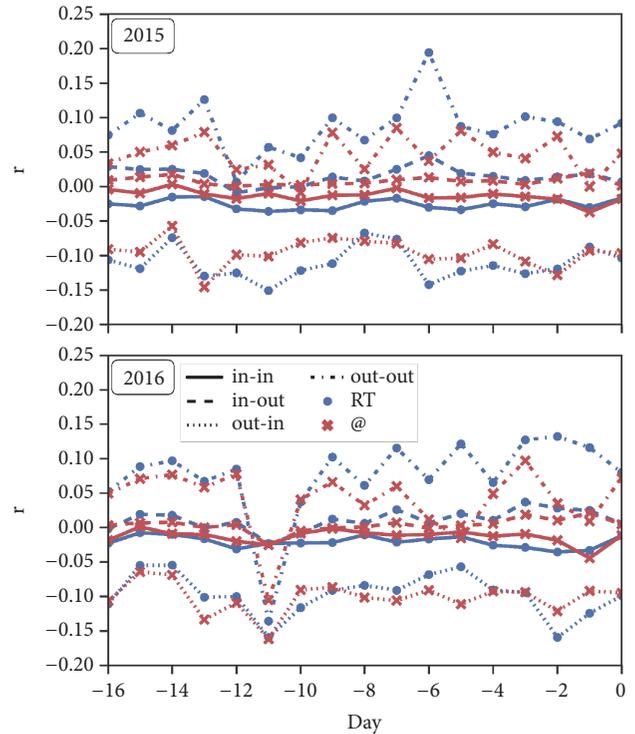


FIGURE 7: Daily evolution of the directed degree assortativities for mentions (@) and retweets (RT) networks for the 2015 (top) and 2016 (bottom) electoral campaigns.

of all the tweets published that day in the analyzed conversation, when the average proportion of retweets received by the most popular tweets each day during the 2016 campaign was $(1.7 \pm 0.3)\%$.

3.3. *Communication Efficiency.* In order to measure the global influence of a user on the network, we have used the user efficiency metric [29]. In Twitter, it can be considered that a message has reached a higher impact if it gets a high number of retweets (RTs). The higher her average number of RTs per original tweet is the more efficient the user is. Hence, the efficiency of a user is defined as follows:

$$\eta_i = \frac{RTs}{Posts} = \frac{(s_{in}^{RT})_i}{A_i} \quad (5)$$

where $(s_{in}^{RT})_i$ is the in-strength of the user i in the retweet network and A_i is her activity, that is, the total number

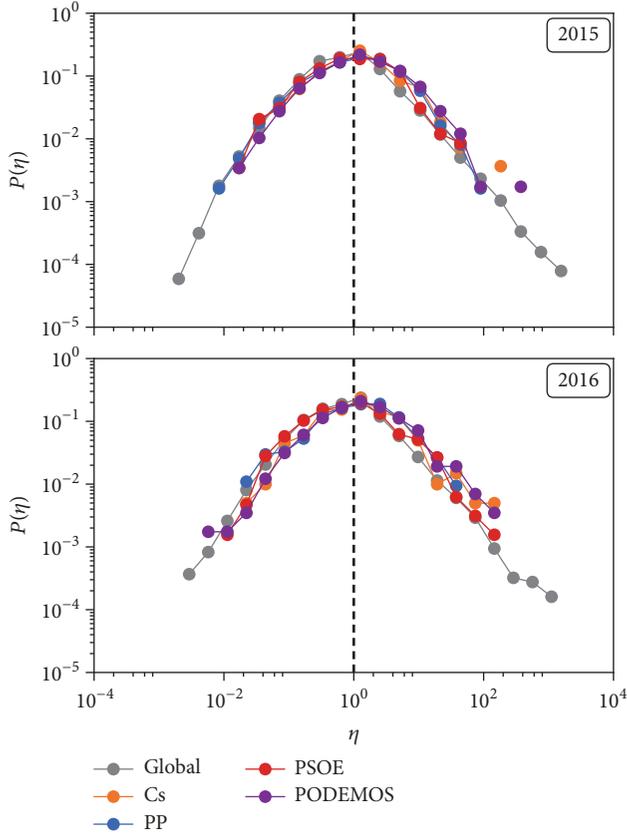


FIGURE 8: Probability distributions of efficiency for Twitter accounts associated with each party and for the whole set of users. Top: 2015 electoral campaign. Bottom: 2016 electoral campaign. The value $\eta = 1$ is marked with a dashed line.

of posted messages. It should be noted that $\eta \in [0, \infty)$. Therefore, this measure represents the collective response to the individual action.

We have computed the distribution of efficiency for the global conversation including regular users and the politicians that participated in it. In order to compare the behavior of regular users with politicians, we have also considered the distribution of efficiency corresponding to the groups of accounts of the four main political parties described in Section 2.2.

In Figure 8 we represent, for both electoral campaigns, the probability distributions of efficiency for Twitter accounts associated with each party and for the whole set of users. This distribution corresponds to the probability mass function. There, we can appreciate that the distributions follow the same functional form for common users and users associated with political parties. In particular, they show a heterogeneous behavior characterized by a heavy tail with power law decay. This has been tested by fitting the tail of the distribution to different functions (exponential, gamma, lognormal, and power law) and choosing the best one according to the Akaike Information Criterion (AIC) [37]. The transition to the power law happens around $\eta \approx 1$. The same kind of behavior is observed in both elections. Notice that the shape of the efficiency distribution is not specific of our electoral context,

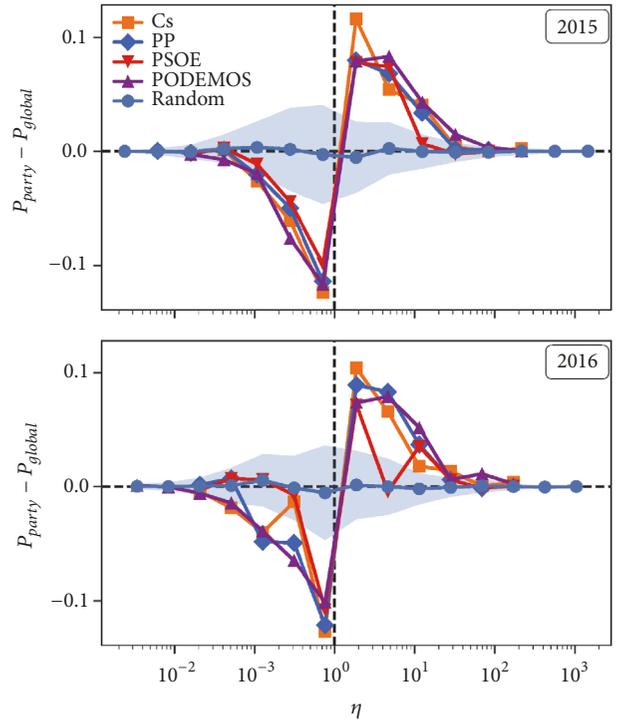


FIGURE 9: Probability differences of having an efficiency that falls within a given bin ($\eta_i < \eta < \eta_{i+1}$) between political users with respect to the total of users. The blue dots and shadowed area corresponds to the average and standard deviation of the differences for 100 samples of 1000 randomly chosen users. Political accounts tend to exhibit higher probabilities of having high efficiencies and lower probabilities of having low efficiencies.

since this structure has been observed in efficiency distributions corresponding to Twitter conversations of different natures and sizes [29].

We have compared the efficiency patterns of the political accounts of each party to the whole set of users for both elections. In order to do that, we have divided the interval that spans all the possible values of efficiency (shown in Figure 8) in logarithmic bins. Then, we have computed the probability of finding an account of a given party in each bin (P_{party}) and the probability of finding a user account belonging to the global conversation (P_{global}). Finally, we have plotted the difference $P_{party} - P_{global}$ for the four main parties as shown in Figure 9.

In order to assess the significance of this result, we have taken 100 samples of 1000 randomly chosen users and computed their efficiency probability differences with respect to the bulk of users as described above. The average values of the differences and their standard deviations are represented, respectively, as blue dots and a blue-grey shadow in Figure 9.

In that figure, it can be noticed that accounts belonging to political parties tend to exhibit higher probabilities than regular users for efficiencies in the region $\eta > 1$ and lower probabilities for efficiencies in the region $\eta < 1$. The behavior of this measure is similar for all the parties. According to this result, the collective reactions to the communication

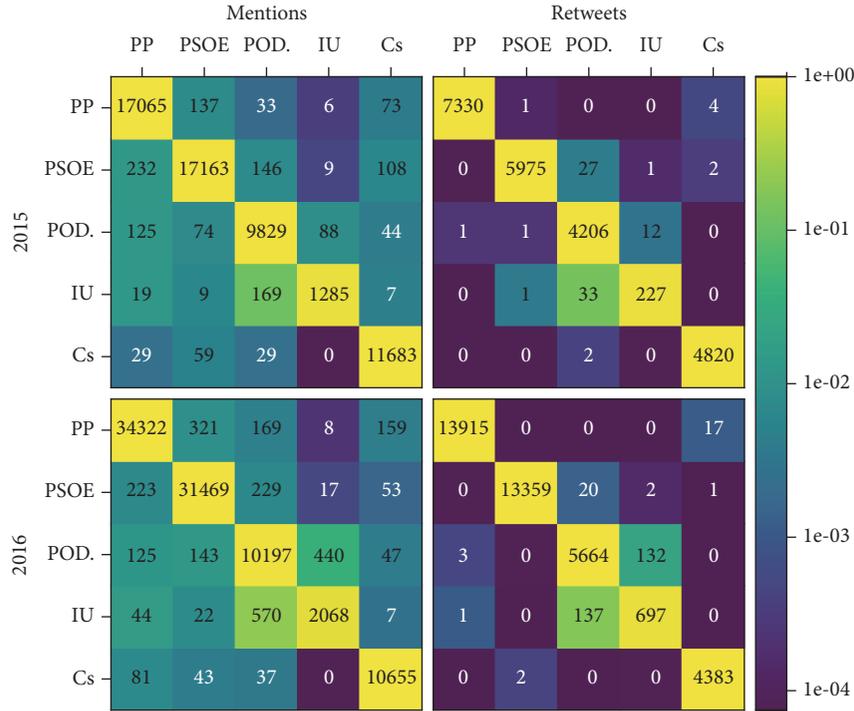


FIGURE 10: Adjacency matrices of C-networks where each supernode is the aggregation of political accounts belonging to a given party. Left panels correspond to mention networks and right panels to retweet networks. Top panels correspond to results of 2015 and bottom panels to results of 2016. Color is related to the proportion of mentions (retweets) directed from party i (row) to party j (column) by the following formula: $c_{ij} = A_{ij} / \sum_j A_{ij}$.

strategies of each party are comparable. Additionally, it is shown that, in general, political accounts are more efficient in propagating messages than regular users.

3.4. Communication among the Elites. While most of the users act as passive listeners or broadcasters, Twitter conversations are usually driven by a small elite of influential accounts [5, 38]. The composition of such elite varies depending on the main subject of the conversation and on the interaction medium [33]. As we have shown in the previous section, politicians are among the most efficient users, implying that they hold a high influence. Given the relevance of this group of users, in this section we will study the communication dynamics among them. To this end, we have retrieved lists of user accounts associated with each party following the process described in Section 2.2.

In order to analyze the communication among politicians, we have computed the subgraphs induced by the user accounts associated with political parties. Then, we have grouped nodes belonging to the same party in supernodes, obtaining a C-network. The resulting supernodes correspond to groups of users and the weights of the links among those supernodes are the sum of the weights of all the links that join a user from one group with a user of another in the original network. The resulting colored adjacency matrices are displayed in Figure 10, where the numbers are the weights of the links, that is, the total number of times that users of a party in a row have mentioned (retweeted) a user of another

party in a column. The color is related to the proportion of mentions (retweets) from party i to party j relative to the total number of mentions (retweets) by party i , such that if A_{ij} is an element of the adjacency matrix, the color of the cell would be proportional to $c_{ij} = A_{ij} / \sum_j A_{ij}$.

As we can see, in line with previous works [5, 15], there is a complete lack of communication between accounts of different parties. The diagonal of the adjacency matrix (the self-links) holds the heaviest weights by more than two orders of magnitude. For the retweet networks, there are almost no links at all between different parties. Taking into account the fact that retweeting a message normally implies an endorsement of the ideas of the original poster, this is not surprising. In the case of the mention networks, there are more message exchanges than in the retweet network.

We have included the party IU in this analysis to study the effect of the agreement between them and Podemos to form a coalition (UP) in the 2016 election. The results displayed in Figure 10 show that, although both parties were already interacting in 2015, they exchanged much more messages in 2016.

We have also computed the evolution of the assortative mixing [39] by political affiliation using two different partitions of the nodes. In the first partition we assign each node to its own party such that we have five groups of nodes corresponding to PP, PSOE, Cs, Podemos, and IU. This is the *partition by party*. In the second partition, we assign the nodes belonging to IU and Podemos to the same group,

leaving the rest of the nodes in their own parties. Hence, we have four groups that correspond to PP, PSOE, Cs, and UP = (Podemos+IU). We have called this second partition, the *coalition partition*.

The assortative mixing is a metric used to test if the links of a network preferentially join nodes of the same kind and nodes of different kinds or the connections are random. In our case, the different kinds would be the different parties in one case and the coalitions in the other. In order to compute the assortative mixing, the nodes are classified in groups and the proportions e_{ij} of strength (that is, the total weights of the links) from group i to group j are computed. In our case, since we already have aggregated the nodes by party in the C-networks, we can use the elements of the adjacency matrix A_{ij} of such C-networks to determine e_{ij} :

$$e_{ij} = \frac{A_{ij}}{\sum_{ij} A_{ij}} \quad (6)$$

Then, $a_i = \sum_j e_{ij}$ is the proportion of links with an origin node of type i and $b_j = \sum_i e_{ij}$ the proportion of links with a target node of type j . If we ignore the network structure, the probability of finding a link with origin node of type i and target node of type j would be $a_i \times b_j$. Taking that into account, the assortative mixing of a network is defined as follows:

$$r_m = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \quad (7)$$

This metric takes the value $r_m = 1$ for a perfectly assortative network, $r_m = 0$ for a random network, and $r_m^{min} = -(\sum_i a_i b_i)/(1 - \sum_i a_i b_i)$ for a perfectly disassortative one.

In Figure 11 we have plotted the temporal evolution of the assortative mixing by party and coalition. As we can see, the assortative mixing by party and coalition remains very high for mentions and retweets. The retweet network is however the most assortative, often reaching values of 1.0.

The drop in assortativity the day after the elections means that politicians of different parties interacted more with each other that day than during the campaign. This seems to be caused partially by some exchange of messages commenting the consequences of the results of the elections. There are tweets containing criticism to adversaries, congratulation messages to related parties, and tweets trying to convince or push potential allies to form coalitions. Notice, however, that this decrease, although significant, is not large: the value reached in 2015 is around 0.87 and in 2016 is 0.89.

Consequently, we attribute the decrease both to the drop in the number of posted messages the day after the elections, which makes the data more noisy, and to an increment of message exchange between parties.

The most remarkable feature of these time series is the difference between the elections of 2015 and 2016. In the latter, the parties IU and Podemos formed a coalition, a fact that is reflected here in the following way: whereas in 2015 coalition and party assortativities are almost equal, in 2016 the coalition assortativity is clearly higher for both networks. This means that the communication between users from Podemos and IU is high enough to lower the general assortativity.

TABLE 6: P values of the paired t-test performed in the assortative mixing time series of parties and coalitions. The test has been carried out by taking each time series of 2015 and coupling it with its counterpart of 2016.

P-value	Parties	Coalitions
Mentions	0.0029	0.9
Retweets	0.0001	0.3

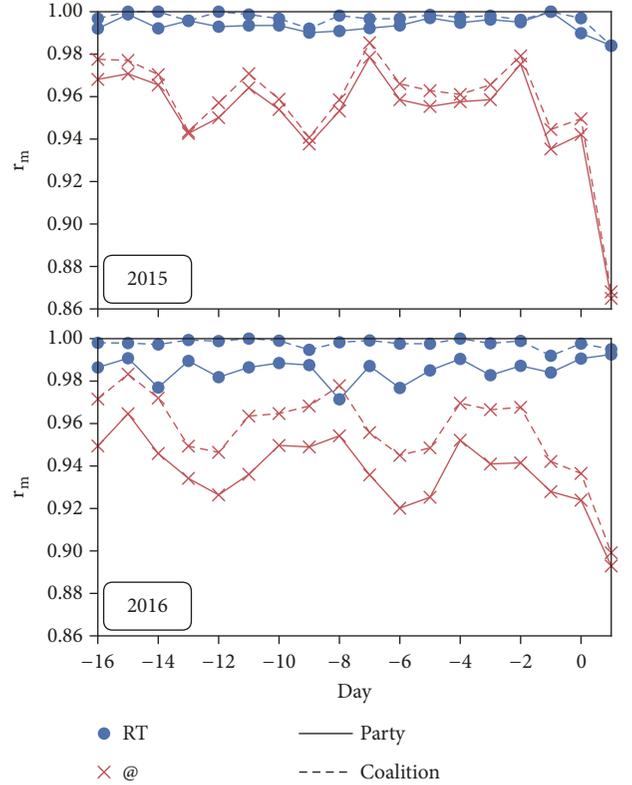


FIGURE 11: Daily evolution of the assortative mixing by party and coalition for the mention (@) and retweet (RT) networks in 2015 (top) and 2016 (bottom) electoral campaigns.

In order to assess the relevance of this effect, we have performed a paired t-test on the assortativities time series coupling each assortativity from 2015 with its counterpart of 2016. The null hypothesis is that the assortative mixing values for both elections have the same expected values. The results presented in Table 6 imply, for a statistical confidence of 99%, a clear rejection of the null hypothesis for the parties time series while in the case of the coalitions time series the null hypothesis can not be rejected.

4. Conclusions

Our main goal in this work was to perform a comparative analysis of the user behavior in Twitter in two consecutive electoral campaigns in order to find the presence of correlations and recurrent patterns. To this end, we have analyzed temporal series and interaction networks corresponding to two Twitter datasets downloaded during the Spanish electoral

campaigns of 2015 and 2016. Although the individual activity and the political actors may have changed, we have shown evidence of recurrent activity patterns in different political campaigns. In particular, the activity time series for both elections exhibit a significant correlation. Moreover, we have found a power law relationship between the daily rate of tweets (retweets and mentions) and the number of unique users. Finally, besides the behavioral stabilities mentioned above, we have been able to detect the effect of a political coalition in the interaction networks through the study of the evolution of their properties.

The results that we have obtained from the computation and analysis of the daily user activity time series for both elections indicate that they present a significant linear correlation. Additionally, by studying the distribution of user activity we have found that in both elections its exponent fluctuates in the same tight interval. The value of the exponent obtained in a previous work [5] also lies within this interval. These facts suggest the existence of recurrent activity patterns in different political campaigns.

We have shown that the daily rate of tweets, retweets, and mentions follow a power law with respect to the number of unique users that participated in the conversation each day. However, whereas in 2015 the growth for the three quantities was slightly super-linear with respect to the number of users, in 2016 we observe an approximately linear behavior. Hence, in 2015, when more users join the conversation, the activity experiences a proportionally higher increment than in 2016.

We have assessed the consistency of the topology of the mentions and retweets networks from one election to the other by computing the degree distribution and the degree correlations of the aggregated networks. The variation of the power law exponent of the degree distributions from one electoral period to the other is of 1% at most, whereas the degree correlations are shifted less than 10% from one year to the other. The values of these properties are also comparable to the results obtained in a previous work with a similar political context [5]. This indicates that the underlying interaction dynamics are robust in the face of a change in social context. However, the analysis of the daily evolution of the degree assortativity of the networks has enabled us to detect fluctuations that seem to be caused by disruptive events, both exogenous and endogenous, and have a relevant impact on the political conversation.

By computing the distribution of the user efficiency for regular users and the accounts associated with each party, we have shown that its functional form is not dependent on the chosen group of users or on the particular electoral period under study. This adds further evidence of the universality of the efficiency patterns shown by Morales et al. [29], where conversations of different natures and sizes were analyzed. We have also computed the differences between users belonging to the global conversation and politicians and found that, in the case of the latter, high efficiencies have a higher probability with respect to regular users whereas low efficiencies present lower probabilities. The behavior of this measure is similar for all the parties. According to this result, the collective reactions to the communication strategies of each party are comparable. Additionally, it is shown that, in

general, political accounts are more efficient in propagating messages than regular users. Politicians are aware of the relevance of social media and know how to leverage their power.

The performed analysis of the mention and retweet C-networks induced by political accounts has enabled us to show the lack of debate among different political parties. This result is in good agreement with the existing literature [5, 15]. Furthermore, we have found that an intensification of the interaction can be detected between parties after the formation of a coalition.

In addition to the regularities in behavioral patterns that we have found by comparing two similar political contexts, several results are consistent with a previous study of the 2011 Spanish elections [5], suggesting that there exist collective behaviors that are robust in the face of social change and can be associated with the Spanish political landscape, but with a potential application beyond this social context. One possible use of these results would be to probe similar political processes and highlight anomalous behaviors that may indicate atypical Twitter uses in electoral contexts.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Disclosure

An earlier version of this work has been presented at the 9th International Conference on Complex Systems.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work has been supported by the Spanish Ministry of Economy and Competitiveness (MINECO) under Contract no. MTM2015-63914-P.

References

- [1] Statista, "Number of monthly active twitter users worldwide from 1st quarter 2010 to 2nd quarter 2018 (in millions)," 2015, <https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>.
- [2] A. Llorente, M. Garcia-Herranz, M. Cebrian, E. Moro, and Y. Moreno, "Social Media Fingerprints of Unemployment," *PLoS ONE*, vol. 10, no. 5, pp. 1–13, 2015.
- [3] S. Asur and B. A. Huberman, "Predicting the future with social media," in *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT 10*, pp. 492–499, IEEE Computer Society, Washington, DC, USA, 2010.
- [4] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welp, *Predicting elections with twitter: What 140 characters reveal about political sentiment*, 2010.

- [5] J. Borondo, A. J. Morales, J. C. Losada, and R. M. Benito, "Characterizing and modeling an electoral campaign in the context of Twitter: 2011 Spanish Presidential election as a case study," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 22, no. 2, p. 023138, 2012.
- [6] B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti, "Geo-located Twitter as proxy for global mobility patterns," *Cartography and Geographic Information Science*, vol. 41, no. 3, pp. 260–271, 2014.
- [7] M. Conover, J. Ratkiewicz, M. R. Francisco, B. Gonçalves, F. Menczer, and A. Flammini, "Political polarization on twitter," in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, vol. 133, pp. 89–96, 2011.
- [8] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito, "Measuring political polarization: Twitter shows the two sides of Venezuela," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 3, p. 033114, 2015.
- [9] A. Herdağdelen, W. Zuo, A. Gard-Murray, and Y. Bar-Yam, "An exploration of social identity: The geography and politics of news-sharing communities in twitter," *Complexity*, vol. 19, no. 2, pp. 10–20, 2013.
- [10] J. H. Parmelee, "Political journalists and Twitter: Influences on norms and practices," *Journal of Media Practice*, vol. 14, no. 4, pp. 291–305, 2013.
- [11] L. C. Abrams and R. Craig Lefebvre, "Obama's Wired Campaign: Lessons for Public Health Communication," *Journal of Health Communication*, vol. 14, no. 5, pp. 415–423, 2009.
- [12] F. I. Solop, "RT @Barack Obama we just made history: Twitter and the 2008 Presidential Election," in *Communicator-in-Chief: How Barack Obama Used New Media Technology to Win the White House*, J. A. Hendricks and J. R. E. Denton, Eds., pp. 37–49, Plymouth: Lexicon Books, 2010.
- [13] G. Caldarelli, A. Chessa, F. Pammolli et al., "A Multi-Level Geographical Study of Italian Political Elections from Twitter Data," *PLoS ONE*, vol. 9, no. 5, p. e95809, 2014.
- [14] D. Cherepnalkoski, A. Karpf, I. Mozetič, M. Grčar, and L. A. Braunstein, "Cohesion and Coalition Formation in the European Parliament: Roll-Call Votes and Twitter Activities," *PLoS ONE*, vol. 11, no. 11, p. e0166586, 2016.
- [15] U. Yaqub, S. A. Chun, V. Atluri, and J. Vaidya, "Analysis of political discourse on twitter in the context of the 2016 US presidential elections," *Government Information Quarterly*, vol. 34, no. 4, pp. 613–626, 2017.
- [16] D. Antonakaki, D. Spiliotopoulos, C. V. Samaras, S. Ioannidis, and P. Fragopoulou, "Investigating the complete corpus of referendum and elections tweets," in *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016*, pp. 100–105, USA, August 2016.
- [17] S. Martin-Gutierrez, J. C. Losada, and R. M. Benito, "Semi-Automatic Training Set Construction for Supervised Sentiment Analysis in Political Contexts," in *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 715–720, Barcelona, 2018.
- [18] D. Gayo-Avello, "A Meta-Analysis of State-of-the-Art Electoral Prediction From Twitter Data," *Social Science Computer Review*, vol. 31, no. 6, pp. 649–679, 2013.
- [19] A. Jungherr, "Twitter use in election campaigns: A systematic literature review," *Journal of Information Technology & Politics*, vol. 13, no. 1, pp. 72–91, 2016.
- [20] V. D. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, Article ID P10008, 2008.
- [21] V. Rodriguez and J. F. OCallaghan, "Spain-government and society," *Encyclopædia Britannica*, 2018, <https://www.britannica.com/place/Spain/Government-and-society>.
- [22] L. Orriols and G. Cordero, "The Breakdown of the Spanish Two-Party System: The Upsurge of Podemos and Ciudadanos in the 2015 General Election," *South European Society and Politics*, vol. 21, no. 4, pp. 469–492, 2016.
- [23] Wikipedia Contributors, *Opinion polling for the spanish general election, 2015*, 2016, https://en.wikipedia.org/wiki/Opinion_polling_for_the_Spanish_general_election,_2015.
- [24] A. J. Morales, V. Vavilala, R. M. Benito, and Y. Bar-Yam, "Global patterns of synchronization in human communications," *Journal of the Royal Society Interface*, vol. 14, no. 128, 2017.
- [25] A. O. Larsson and H. Moe, "Studying political microblogging: Twitter users in the 2010 swedish election campaign," *New Media & Society*, vol. 14, no. 5, pp. 729–747, 2012.
- [26] A. Jungherr, "Tweets and votes, a special relationship: The 2009 federal election in germany," in *Proceedings of the 2nd Workshop on Politics, Elections and Data*, pp. 5–14, ACM, New York, NY, USA, 2013.
- [27] Organic Law 5/1985, of the 19 of June, on the general electoral system, BOE núm, 1985, pp. 19110–19134, <https://www.boe.es/boe/dias/1985/06/20/pdfs/A19110-19134.pdf>.
- [28] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graphs over time: densification laws, shrinking diameters and possible explanations," in *Proceeding of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '05)*, pp. 177–187, New York, NY, USA, August 2005.
- [29] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito, "Efficiency of human activity on information spreading on Twitter," *Social Networks*, vol. 39, no. 1, pp. 1–11, 2014.
- [30] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661–703, 2009.
- [31] T. P. Peixoto, "Hierarchical Block Structures and High-Resolution Model Selection in Large Networks," *Physical Review X*, vol. 4, no. 1, p. 011047, 2014.
- [32] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.
- [33] J. Borondo, A. J. Morales, R. M. Benito, and J. C. Losada, "Multiple leaders on a multilayer social media," *Chaos, Solitons & Fractals*, vol. 72, pp. 90–98, 2015.
- [34] J. G. Foster, D. V. Foster, P. Grassberger, and M. Paczuski, "Edge direction and the structure of networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 24, pp. 10815–10820, 2010.
- [35] M. E. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 64, no. 2, 2001.
- [36] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX," in *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pp. 11–15, Pasadena, CA, USA, 2008.
- [37] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, 1974.

- [38] J. He, L. Hong, V. Frias-Martinez, and P. Torrens, "Uncovering Social Media Reaction Pattern to Protest Events: A Spatiotemporal Dynamics Perspective of Ferguson Unrest," in *Social Informatics*, vol. 9471 of *Lecture Notes in Computer Science*, pp. 67–81, Springer International Publishing, Cham, 2015.
- [39] M. E. Newman, "Mixing patterns in networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 67, no. 2, p. 026126, 2003.

