The new Spanish political scenario: Twitter graph and opinion analysis with an interactive visualisation

Pelayo Quirós*
Blanca Rosario Campomanes-Álvarez*
pelayo.quiros@ctic.es
charo.campomanes@ctic.es
CTIC Technological Centre
Gijón, Asturias, Spain

ABSTRACT

Twitter, as a social network, has been used extensively as a way to express users opinion about a wide range of topics. Particularly, politics have been one of the main issues in the last years, and consequently, its study in Twitter can lead to valuable conclusions.

The methodology presented in this paper is based on the extraction of the followers network of the main politicians in Spain prior to the general election of November of 2019, providing a suitable graph analysis and visualisation. Hence, closeness between different political parties can be spotted with respect to their common followers. This is motivated by the new political scenario, given that in the last five years, it derived from a bipartisan system to a plural one, with six relevant national parties.

Additionally, information about each political party is presented interactively with respect to the political profile of those who write about them. Reciprocally, this tool also provides information about the target of this party identified profiles.

KEYWORDS

Twitter, political analysis, social networks, graphs, data visualisation

1 INTRODUCTION

Twitter is a microblogging social network based on the publication of short messages of 280 characters at most. It has been used extensively as an opinion network with daily volumes of generated messages of about 500 million tweets, with 139 million daily active users (https://business.twitter.com/). Every user has the option to follow other accounts, so the content published by them is shown to their followers. The different actions with respect to each Twitter message or tweet are sharing it with your followers (retweet), marking it as favourite or answering the tweet. Furthermore, every individual can mention another user by using the account name preceded by an "@".

Consequently, many approaches have been studied concerning different types of insights. Sentiment analysis has been one of the most relevant ([29], [22], [21], [1], [14]), as well as graph analytic of the connection network among Twitter users ([11], [27], [26]) or of the graph generated by their communication through mentions, answers and retweets ([34], [8], [37]).

Furthermore, these applications are usually centred in a particular field, given the raw daily volume, as it has been previously mentioned. Among others, economics ([33], [25], [10]), television and cinema ([2], [12], [30]), or sports ([28], [36], [31]) have been widely studied. However, this paper is focused on the application of Twitter analytic to politics, which has already been tackled as well ([15], [16], [4], [5]). Particularly, this analysis is devoted to the Spanish political situation in Twitter prior to the second general election in 2019, which took place the 10$^{\text{th}}$ of November.

The main motivation of this approach is based on the recent evolution of the Spanish political scenario (http://www.infoelectoral.mir.es/). It has been a two-party system since the general election of 1989, where PSOE and PP received most of the votes. Particularly, the total percentage of votes obtained by these two parties ranged from 65% in 1989, to 73% in 2011, getting their highest combined percentage in 2008 with a total over the 83% of votes. However, in 2015 two new parties irrupted, Unidas Podemos and Ciudadanos, reducing dramatically the bipartisan percentage to 50% of votes. Another two new political parties were added to this scenario in 2019, VOX and Más País, turning the sum of votes of PSOE and PP, for the first time since their coexistence, into values under the 50% (45% in April, 48% in November). Additionally, regionalist parties have strengthen their presence in their respective areas, with a sum of 10% and 11% of the total votes in both general elections in 2019, respectively.

As a consequence, with such a variety of political parties, it is hard to assess the closeness between certain parties. Even though the leaders may point out their preferences, their voters may have a different opinion, and as such, analysing their closeness with respect to their supporters is a must in order to understand the underlying structure of the actual political Spanish scenario.

To analyse such closeness with respect to supporters, all the followers of the most remarkable politicians associated to the ten most relevant political parties have been downloaded via Twitter API, including the two bipartisan parties (PSOE, PP), the national ones of recent creation (Unidas Podemos, Ciudadanos, VOX, Más País), and the four most relevant regionalist parties (ERC, Junts Per Catalunya, PNV, EH Bildu).

All this information has been processed in order to put it together as a graph or network. The main goal of this work is to better understand the relationship between user and political Twitter accounts, as well as discovering insights about the new and uncertain political Spanish scenario. This could be done by means of Social Network Analysis (SNA) techniques due to the nature of the used data.

SNA is used for measuring and analyzing the structural properties of networks of interdependent dyadic relationships. One of the core assumptions of SNA is that the patterns of these relationships can have important effects on individual and organizational

*Both authors contributed equally to this research.

behaviour, constraining or enabling access to resources, and exposure to information and behaviour [17]. Apart from this, one of the key elements that characterizes modern SNA is the use of visualisations of complex networks. Innovators in information visualisation have also contributed to helping users to discover patterns, trends, clusters, gaps, and outliers, even in complex social networks [32].

In order to complement this study, an analysis of the different messages directed to each of the candidates in Twitter has been developed, obtaining the political typology of those who speak about them.

This paper is structured as follows. Section 2 presents the whole proposed methodology. In Section 3 the development of the application, along with visual examples, is provided. Section 4 is devoted to conclusions and future work.

2 PROPOSED METHODOLOGY

As previously said, the recently appeared new parties in Spain have added important and unknown changes to the political tendencies in the public opinion. This constitutes a new scenario in which voters have changed their preferences from only two options to a system with multiple parties, with a significant rise of nationalist ones. Understanding and visualising these new political trends and how the vote-flow can vary from one party to another or how parties are related between them, arises as an interesting task for interpreting the new complex political scenario in Spain.

The tool presented here is based on the study of the large network that represents the different connections among Twitter users, where a relationship between two given individuals is not necessarily reciprocal, which leads to a directed graph study corresponding to what users follow each individual.

Particularly, this work has been devoted to analysing the Spanish political Twitter environment, previous to the general election that took place the 10th of November of 2019. The ten main political parties in Spain have been monitored: PSOE, PP, VOX, Unidas Podemos, Ciudadanos, Más País, ERC, Junts Per Catalunya, PNV and EH Bildu. The first six are national parties, while the remaining ones are only present in certain areas (Catalonia for ERC and Junts Per Catalunya, Basque Country for PNV and EH Bildu).

Hence, the proposed approach aims to perform an interactive network by: 1) extracting and processing data based on political interactions from Twitter, 2) from this data, creating a graph and extracting communities from it, 3) applying a layout algorithm to align nodes according to their community membership; and 4) producing a two-dimensional graph visualisation in a web-based platform that supports pan, zoom navigation and an additional data analysis of the relevant nodes by means of the interaction with them.

The whole architecture of the proposed system is depicted in Figure 1. This diagram shows two connected lines of action. The first one is based on the extraction of followers of each politician from the ten political parties, which are afterwards clustered and filtered, obtaining a set of followers reduced and processed, in order to apply the different graph methodologies with the open-source SNA software Gephi (version 9.1) [6]. Furthermore, the original followers dataset is used to obtain political profiles with respect to that ten parties.

Secondly, a daily download of tweets that mention the six national political parties and their respective leaders is conducted. These messages are processed to obtain the profile of opinion

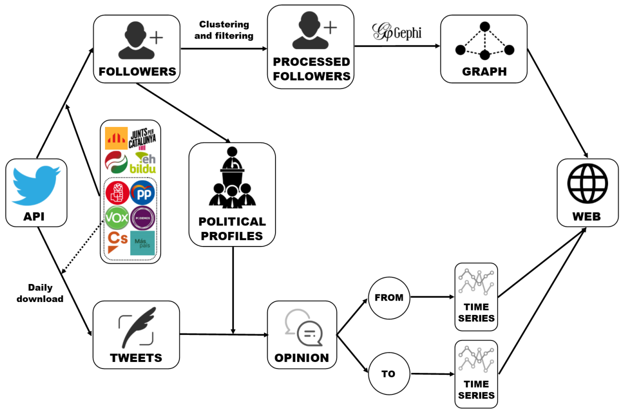

Figure 1: Diagram of the proposed methodology

from and to the selected party, with the previously obtained profiling set. Both graphical outputs are presented in a website that allows a joint visualisation. This web application is based on the Node.js [13] engine. It has been developed in order to carry out a post visual data exploration and an interactive analysis of the network.

2.1 Data Extraction and Processing

Data extraction is carried out by using the official Twitter API (https://developer.twitter.com/), which allows to obtain a wide range of information from Twitter, such as timelines, tweets mentioning a certain term, hashtag or account, or the users that follow a given individual. In this paper, the proposed approach is based on the latter one. Consequently, it is mandatory to decide which accounts should be analysed for every political party. For each one, the official account from the party, their corresponding leaders, as well as additional remarkable politicians have been monitored. In Table 1 a list with each political party and the number of accounts analysed for each one is presented.

Table 1: Summary of accounts analysed for each political party

Political party	Number of accounts
PSOE	11
PP	11
VOX	10
Unidas Podemos	11
Ciudadanos	11
Más País	11
ERC	9
Junts Per Catalunya	8
PNV	7
EH Bildu	7
TOTAL	96

As it can be observed, the number of accounts differs slightly among them. This is due to the lack of relevant and visible additional politicians for the minor parties. Globally, 96 political accounts have been selected.

On the other hand, a network consists of two components, a list of the actors which compose the network, and a list of relations, i.e., the interactions between actors. As part of a mathematical object, actors will then be called vertices or nodes, and relations will be denoted as edges. Based on graph theory, a network corresponds to a graph, which is built by vertices or nodes as well as edges connecting vertices between them. This way of representing data is appropriate for scenarios like the proposed one, which involves a high number of connections and can not be accurately understood and represented by using traditional graphics due to its complex structure.

Taking into account the extracted data and the graph theory, a directed network can be built from the generated dataset. Two types of nodes are defined. The first type of them refers to the political accounts, whereas the other one determines the user accounts, i.e., individuals following political accounts. Regarding the interaction between nodes, an edge represents the relationship that is created when a user follows a political account.

Once the initial information is downloaded, as it has been previously explained, it is necessary to store and process it to apply the different graph processing, visualisation and analysis techniques.

Given the complexity granted by data dimension, it is mandatory to simplify it, so it can be manageable by usual graph tools. In order to do so, the original data is modified by aggregating every user who follows exactly the same political accounts into a common cluster, whose node size is associated to the number of individual that it represents. Furthermore, each node has an edge to each political account that this cluster of accounts follows.

Simultaneously, it is developed a methodology that determines which users are really interested in one or many political parties, while applying a filter that removes those whose interest is weak, circumstantial or too spread among several political parties. This information is complementary used to the graph-associated one, in order to enrich the provided visualisation.

2.2 Graph Processing: Community Detection Algorithms

Community detection algorithms in graphs are used for identifying groups of similar individuals in order to understand user interactions and behaviours. In addition, some of these behaviours are only observable into a group and not on an individual level. This is because individual behaviours could easily change, but collective behaviours are more robust to changes [7]. These algorithms are also useful to define the communities of highly related nodes as well as visualising their relations to other communities.

In this work, the algorithms considered for identifying different communities in the graph, are the following:

- The Louvain method [9]: This algorithm detects communities in networks by maximizing a modularity score for each community, where the modularity quantifies the quality of an assignment of nodes to communities by evaluating how much more densely connected the nodes within a community are, compared to how connected they would be in a random network. This method is one of the fastest modularity-based algorithms, and works well with large graphs. It also reveals a hierarchy of communities at different scales, which can be useful for understanding the global functioning of a network.
- The Leiden algorithm [35]: This algorithm can be seen as an improvement of the Louvain algorithm. The Leiden algorithm also takes advantage of the idea of speeding up the local movement of nodes and the idea of moving nodes to random neighbours. It consists of three phases: (1) local movement of nodes, (2) refinement of the partition and (3) aggregation of the network based on the refined partition, using the non-refined partition to create an initial partition for the aggregate network.

Once the community detection is performed, the next step consists of evaluating the results obtained by the community-detection algorithms. The selected metric for determining the quality of a community is the modularity [24]. This measure is based on the idea that a random graph is not expected to have a cluster structure. Therefore, the possible existence of clusters is revealed by the comparison between the actual density of edges in a subgraph and the density that would be expected in a random subgraph.

Furthermore, other centrality measures can be calculated in order to provide the importance, or influence, in a social network. For instance, the Degree Centrality indicates what accounts are the most followed, the Eigenvector Centrality shows which are the most influential accounts, and the Betweenness Centrality detects who/which are the users/accounts controlling the information flow [38].

2.3 Graph visualisation

Once the communities are calculated, a colour for each of them (nodes and edges in a particular community) are assigned to better distinguish the structure of the different groups inside the graph. In addition, the node size is also determined by firstly distinguishing between political and user nodes. The size of the political nodes is modified based on the number of followers they have, i.e., political nodes with a higher in-degree present a larger size than political nodes with a lower in-degree. On the other hand, the size of the user nodes will be larger when these nodes group together more individuals.

Regarding the information of each node, the visualisation of the graph also involves showing the data related to each political node such as the name of the political party or the photograph of the politician. This could be useful, for instance, to rapidly distinguish if nodes of different political parties belongs to the same community or in which position of the political scenario (the calculated graph) they are located.

After establishing the appearance of the whole network, a layout algorithm should be applied for drawing the graph in an aesthetically way as well as differentiating the communities within the graph.

Their purpose is to position the nodes of a graph in a two-dimensional space so that all the edges are of more or less equal length and there are as few crossing edges as possible. This is done by assigning forces among the set of edges and the set of nodes, based on their relative positions, and then using these forces either to simulate the motion of the edges and nodes or to minimize their energy [3].

The considered layout-graph algorithms for modelling the shape of the graph are the following:

- ForceAtlas2 [18]: This algorithm handles large networks while keeping a very good quality. Nodes repulsion is approximated with a simulation, which therefore reduces the algorithm complexity.
- OpenOrd [23]: This algorithm aims to better distinguish clusters via a simulated annealing type schedule. Long edges are cut to allow clusters to separate.

2.4 Analysis and Additional visualisation

In addition to the Gephi [6] analysis of the graph, a web-based platform has been developed with the aim of performing interactions on the graph previously built. This web, based on Node.js [13], provides navigation options such as pan or scroll through the data in order to drill-down and access details of each part of the graph. Selecting and zooming can be used to facilitate quick and interactive exploration of data in order to see connections between nodes and communities. Smooth zooming is used in order to explore a region of interest.

Furthermore, it could be possible to visualise statistics related to a set of connected nodes by selecting a particular node. This analysis is based on the following.

As it has been stated in Subsection 2.1, the initial users associated to the followers of the studied politicians have been analysed also with respect to which political party or parties are they interested in, obtaining a political profile for more than 4 million of users for the ten selected parties. Furthermore, an automatic daily process has been designed to download every tweet that mentions the leaders and official accounts of the main six political parties (PSOE, PP, VOX, Unidas Podemos, Ciudadanos, Más País). Once this information is processed as well, it is possible to merge both datasets in order to obtain valuable information, being able to analyse, in a daily basis, the profile of the authors of every downloaded message.

Thus, it is possible to analyse, for each party, what are the political profiles that speak the most about it. Inversely, the most common political targets of the messages generated by a given profile can be obtained as well.

This information is generated every day, so it is possible to obtain temporal series associated to each political party with respect to both aspects: the profile of those who speak about them, and the profile of those who they speak about. This eases the visualisation of such information, being able to get a global image and evolution of the opinion involving each party in both directions.

3 APPLICATION DEVELOPMENT

All the experiments have been performed on an Comet Lake i7-10710u CPU 1100 MHz, with 32 GB RAM and a GEFORCE GTX1650 MAX-Q graphic card, running Ubuntu 18.04.3 LTS.

3.1 Data Collection and Processing

The data extraction process using the Twitter API is conducted with the statistical programming language R, which provides a package developed to apply the different functionalities of the Twitter API (rtweet) [20], after obtaining a mandatory API key. Twitter API is free, although it has volume restrictions by period of time, so once this limit is reached, the download is stopped for about 15 minutes. However, there are no restrictions concerning the total downloadable volume, so it is possible to obtain the required information.

The 96 political accounts distributed in ten parties as stated in Table 1, have been studied by extracting all the users that follow each of the these accounts. As a result, a total of 23.305.485 connections (edges of the graph), corresponding to 6.241.682 unique users (nodes of the graph) have been obtained. This information has been stored in a MongoDB database, as well as every additional information generated from this initial dataset. Consequently, two tables have been generated and stored in the database, where the first one describes the different nodes with their respective node sizes, and the second one describes the directed edges that connect these nodes.

As it has been aforementioned, a simplification of this data is designed, grouping followers in clusters with the same pattern with respect to what politicians they follow. This simplification leads to a nodes table with 124.377 entries, and an edges table with 898.309 rows. It should be noted that this simplification does not imply removal of information, as the individual data of each user is not relevant for this paper's purposes. However, this data has many negligible nodes due to their small size, so a filter has been applied, removing those whose node size is less or equal than 100. After that, the size of the nodes and edges tables are considerably reduced to 2.863 and 12.291, respectively. These removed clusters are representative of such a small portion of individuals (less than 100 users in a population of over 6 million), that the vast visualisation improvement makes up for the information loss. This process allows a better visualisation, as such a large number of nodes and edges would lead to a hard to interpret graph with insignificant noise.

On the other hand, the parallel process to obtain political profiles has been also applied. With this process, 4.567.607 accounts have been identified with real interest in at least one of the studied political parties. In Table 2, a summary of these users is provided, with a list of the total number of individuals that have been identified as interested in each political party.

Table 2: Number of users identified as interested in each political party

Political party	Number of interested users
PSOE	843.226
PP	553.048
VOX	300.569
Unidas Podemos	2.332.736
Ciudadanos	885.340
Más País	521.249
ERC	671.540
Junts Per Catalunya	203.204
PNV	16.908
EH Bildu	41.466

Note that the sum of these values is greater than the previous number of analysed accounts, as many of them show interest in more than one party simultaneously. Particularly, the three most common combinations are PSOE & Unidas Podemos, Unidas Podemos & Más País and ERC & Junts Per Catalunya with 310.527, 187.417 and 161.408 occurrences, respectively. These pairs are coherent, given that the ones that compose the first two pairs are the most voted national left wing political parties, and that ERC and Junts Per Catalunya are both in favour of the independence of Catalonia.

Additionally, as it has been previously explained, tweets mentioning the accounts of the six political national parties and their respective leaders are downloaded daily. This process has been started in January of 2019, and up to November of 2019, the total of downloaded messages is over 15 million tweets, with about 50.000 tweets downloaded every day in average.

3.2 Experimental Design for the Application

First of all, experiments have been carried out in order to compute several centrality measures of the initial graph. In particular, the computed metrics were the average degree, the graph density and the number of connected components. Secondly, the performance of the two community detection algorithms: Louvain [9] and Leiden [35] was tested. The method with the best behaviour in terms of modularity was selected. The selected configuration for these two methods was resolution equal to 0.01, 10 iterations and random seed for initializing the algorithms.

In order to improve the representation of the graph with its corresponding communities, an adjustment in colour and size was done as previously explained in Subsection 2.3. Regarding the configuration, the size values were ranging between 10 and 200, where 10 corresponds to the minimum size and 200 to the maximum one.

Following these tasks, ForceAtlas2 [18] and OpenOrd [23] layout algorithms were consecutively combined to customize the network in order to improve its visualisation. The initial parameters for ForceAtlas2 were shown in Table 3. On the other hand, the OpenOrd algorithm was set with an edge cut of 0.95 (a higher cutting means a more clustered result), a number of iterations of 250 and random seed.

After this, a set of centrality measures was computed for giving insights about the connections and relationships between communities: in-degree, out-degree, Betweeness Centrality, Closeness Centrality and Eigenvector Centrality [19].

Table 3: ForceAtlas2 initial configuration

Parameter	Value
Threads number	8
Tolerance	0.9
Approximate repulsion	yes
Approximation	1.2
Scaling	1.5
Stronger gravity	no
Gravity	1.0
Dissuade hubs	no
LinLog mode	no
Prevent overlap	yes
Edge weight influence	1.0

Finally, the parameters of both the resulting graph and the layout were saved in a JSON file with the aim of exporting it to the web application based on Node.js [13]. This front-end constitutes a web platform for allowing different actions on the graph like visualisation, zoom navigation and additional data analysis of the relevant nodes by means of the interaction with them.

3.3 Results

Regarding the metrics that calculate the network properties as a whole, the obtained results were the following. For the graph density, the value was equal to 0.003. Closely related to the density of the graph is the average degree. The value obtained for this metric was 3.24, i.e., the average number of edges connected to a node. Finally, the number of strongly connected components were nine, being each component a maximal strongly connected subgraph.

Table 4 shows the results achieved by the community detection algorithms. Both methods obtain nine significant communities inside the network, while the Leiden algorithm outperforms the Louvain method in a 0.6%. These results are consistent with the fact that the number of strongly connected components is the same than the number of communities obtained by the tested algorithms.

Table 4: Results of the community detection algorithms

Algorithm	Nº clusters	Modularity
Louvain	9	0.990
Leiden	9	0.996

In particular, the characteristics of the obtained clusters can be summarized as follows. On the one hand, a significant cluster that groups together the regionalist parties from Catalonia with their followers was detected. On the other hand, the algorithm identified two different clusters composed by regionalist parties from Basque Country and followers. The VOX party accounts and its followers, were represented in a fourth community, which appears clearly separated from the other groups. In another cluster, some political accounts from Más País and Unidas Podemos leaders were classified. However, followers and other political accounts of these latter parties were also identified in another community, closer to regionalist parties. Regarding the national parties, the main leaders and political accounts of Ciudadanos and PP were group together. Additionally, the current socialist president and other accounts from PSOE composed the eight cluster. Finally, the algorithm differentiated in a separate and central community, the Spanish ex-president from the PP party as well as the PP and PSOE official Twitter accounts and significant politicians from these two parties.

With respect to the centrality metrics, four different options have been considered: in-degree (number of adjacent incoming edges to each node), Closeness (steps required to access every other node), Betweeness (based on the number of shortest paths between nodes that pass through a particular node) and Eigenvector Centrality (connection to well-connected nodes), These measures have been obtained for every political node that has at least a connection with two clusters. In order to aggregate that information, given the heterogeneous scale, a ranking has been generated for each node, with greater score to the ones with the better values. The sum of such values defined the final rank of these nodes. The most connected nodes are the ones shown in Table 5.

Table 5: Best political nodes with respect to the centrality metrics

Rank	Account	Party	Score
1	@Pablo_Iglesias_	Unidas Podemos	352
2	@Albert_Rivera	Ciudadanos	345
3	@ahorapodemos	Unidas Podemos	342
4	@ManuelaCarmena	Más País	335
5	@sanchezcastejon	PSOE	335
6	@agarzon	Unidas Podemos	334
7	@marianorajoy	PP	332
8	@KRLS	Junts Per Catalunya	326
9	@ierrejon	Más País	322
10	@PSOE	PSOE	314
11	@gabrielrufian	ERC	313
...	...	...	...
20	@vox_es	VOX	270
33	@ArnaldoOtegi	EH Bildu	213
47	@eajpnv	PNV	167

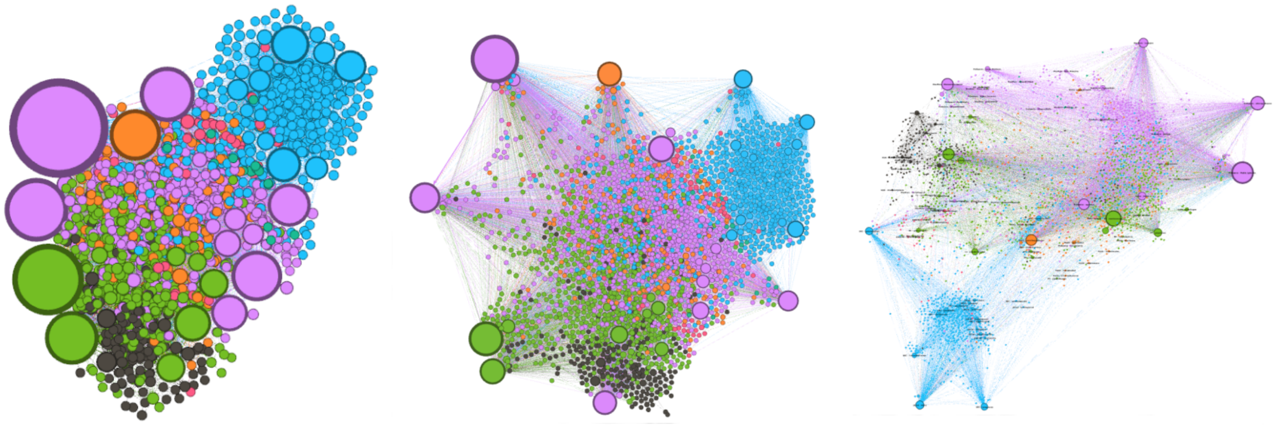

Figure 2: From left to right evolution of graph visualisation: initial and in-progress layout options

These results show that the nodes with the best centrality trade-off are coped by accounts from Unidas Podemos (1^{st}, 3^{rd}, 6^{th}), while parties like VOX, EH Bildu and PNV get their first account in that ranking in the positions 20^{st}, 33^{th} and 47^{th}, respectively.

Additionally, the visualisation of the graph by using ForceAtlas2 [18] and OpenOrd [23] obtained a set of representations depicted in Figure 2. The final layout was selected for being saved and represented on the web platform.

In Figure 3, a screenshot of the most dense part of the graph from the aforementioned web tool is presented. It is possible to spot a group of political nodes corresponding to parties in favour of the independence of Catalonia (ERC, Junts Per Catalunya), as well as the national party VOX isolated from the rest. The other national parties are all together in the centre of the graph, although with more closeness between those with more ideological similarities.

Figure 4 presents the additional graphics that show up once a political node is selected in the interactive graph. The first one provides information about the messages written by the followers of the selected party. The second one represents the share of opinion directed to that party with respect to the political profile of their authors.

4 CONCLUSIONS

A methodology has been proposed to obtain valuable political information from Twitter. Particularly, the Spanish scenario has been considered, taking into account the second general election of 2019, motivated by the changing political situation, where in five years the system went from bipartisan to multi-party with six national parties and several regionalist ones.

This analysis is twofold. Firstly, several political accounts from the ten main political parties have been selected, and their followers have been obtained through the official Twitter API. This information, after the necessary processing, is treated as node and edge data, and as a consequence, a proper graph analysis and visualisation is provided. Secondly, tweets mentioning the official accounts of the main six national parties and their respective leaders are downloaded in a daily basis. These tweets are characterised with respect to the political profile of their authors with the data generated by the followers information. Consequently, it is possible to determine for each party, whose followers are speaking about it, and what party are its followers talking about as well.

All the obtained visual representations have been merged into one in a website designed to that use. Thus, a graph showing the political nodes, the clusters, their connections and the generated communities, is presented interactively. Additionally, once a political node is selected, time series are shown with respect to the opinion from and to that party.

It is possible to draw important conclusions from this analysis, such as the political communities where most accounts from the same party are set together despite the system not knowing the corresponding political party, as well as the closeness between parties that share some common ideology. Additionally, opinion analysis shows the existence of parties whose followers generate a great share of the content directed to this particular party.

Particularly, the resulting graph shows that the Catalonian parties (ERC, Junts Per Catalunya) are in the same cluster taking into account their common ideology regarding Catalonia independence. VOX is also isolated from the others, showing their auto-connection and lack of interaction with the others. The rest of national parties (PSOE, PP, Unidas Podemos, Ciudadanos, Más País) are together, but with visible closeness of those with more ideological similarities.

Future work leads to enrich the additional information provided about each party interactively in the graph along with the opinion time series, using the Twitter profile information of each analysed user. Additionally, new lines of analysis of politics in Twitter are to be considered, which can complement the already developed ones.

REFERENCES

[1] Apoorv Agarwal, Boyi Xie, Ilia Vovsha, Owen Rambow, and Rebecca Passonneau. 2011. Sentiment Analysis of Twitter Data. In *Proceedings of the Workshop on Languages in Social Media (LSM '11)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 30–38. http://dl.acm.org/citation.cfm?id=2021109.2021114

[2] Akshay Amolik, Niketan Jivane, Mahavir Bhandari, and M. Venkatesan. 2016. Twitter Sentiment Analysis of Movie Reviews using Machine Learning Techniques. *International Journal of Engineering and Technology* 7 (01 2016), 2038–2044.

[3] Michael J. Bannister, David Eppstein, Michael T. Goodrich, and Lowell Trott. 2013. Force-Directed Graph Drawing Using Social Gravity and Scaling. In *Graph Drawing*, Walter Didimo and Maurizio Patrignani (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 414–425.

[4] Pablo Barberá. 2015. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. *Political Analysis* 23, 1 (2015), 76–91.

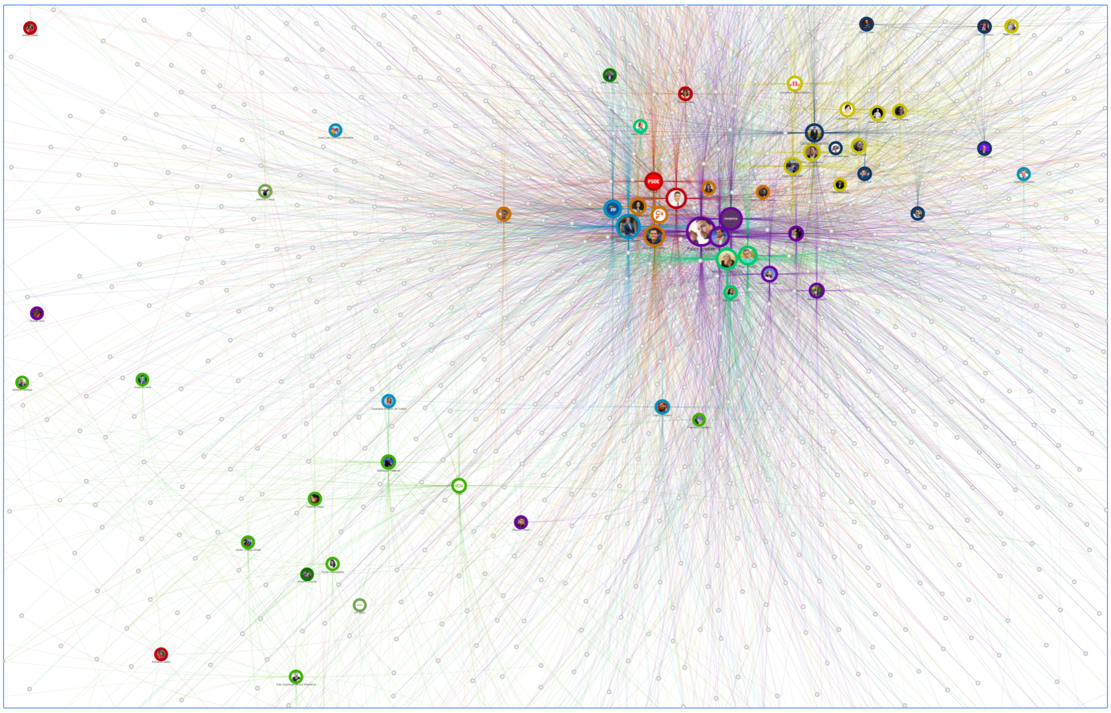

Figure 3: Screenshot of the central part of the graph in the web platform

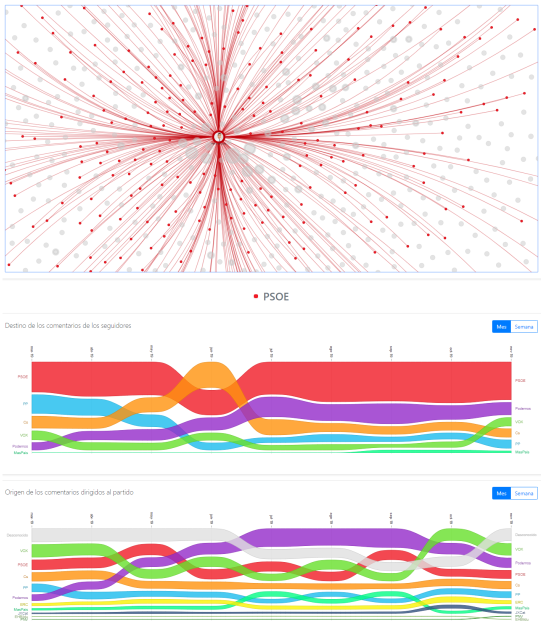

Figure 4: Graph after the selection of a political node (top), with visualisation of the opinion from (middle) and to (bottom) the selected party (PSOE)

https://doi.org/10.1093/pan/mpu011

[5] Pablo Barberá and Thomas Zeitzoff. 2017. The New Public Address System: Why Do World Leaders Adopt Social Media? *International Studies Quarterly* 62, 1 (10 2017), 121–130. https://doi.org/10.1093/isq/sqx047 arXiv:http://oup.prod.sis.lan/isq/article-pdf/62/1/121/24431285/sqx047.pdf

[6] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154

[7] Gema Bello-Orgaz, Julio Hernandez-Castro, and David Camacho. 2017. Detecting discussion communities on vaccination in twitter. *Future Generation Computer Systems* 66 (2017), 125 – 136. https://doi.org/10.1016/j.future.2016.06.032

[8] David R. Bild, Yue Liu, Robert P. Dick, Z. Morley Mao, and Dan S. Wallach. 2015. Aggregate Characterization of User Behavior in Twitter and Analysis of the Retweet Graph. *ACM Trans. Internet Technol.* 15, 1, Article 4 (March 2015), 24 pages. https://doi.org/10.1145/2700060

[9] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008, 10 (2008), P10008. http://stacks.iop.org/1742-5468/2008/i=10/a=P10008

[10] Johan Bollen, Alberto Pepe, and Huina Mao. 2009. Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *CoRR* abs/0911.1583 (2009). arXiv:0911.1583 http://arxiv.org/abs/0911.1583

[11] Meeyoung Cha, Hamed Haddadi, Fabrício Benevenuto, and Krishna P. Gummadi. 2010. Measuring User Influence in Twitter: The Million Follower Fallacy. *AAAI Conference on Weblogs and Social Media* 14.

[12] Alfonso Crisci, Valentina Grasso, Paolo Nesi, Gianni Pantaleo, Irene Paoli, and Imad Zaza. 2018. Predicting TV programme audience by using twitter based metrics. *Multimedia Tools and Applications* 77, 10 (01 May 2018), 12203–12232. https://doi.org/10.1007/s11042-017-4880-x

[13] Ryan Lienhart Dahl. 2009 (accessed 30-November-2019). Node.js Foundation. https://nodejs.org/.

[14] Sahar A. El Rahman, Feddah A. AlOtaibi, and Wejdan. A. AlShehri. 2019. Sentiment Analysis of Twitter Data. In *2019 International Conference on Computer and Information Sciences (ICCIS)*. 1–4. https://doi.org/10.1109/ICCISci.2019.8716464

[15] Tarek Elghazaly, Amal Mahmud, and Hesham Hefny. 2016. Political Sentiment Analysis Using Twitter Data. 1–5. https://doi.org/10.1145/2896387.2896396

[16] Ratab Gull, Umar Shoaib, Saba Rasheed, Washma Abid, and Beenish Zahoor. 2016. Pre Processing of Twitter's Data for Opinion Mining in Political Context. *Procedia Computer Science* 96 (2016), 1560 – 1570. https://doi.org/10.1016/j.procs.2016.08.203 Knowledge-Based and Intelligent Information Engineering Systems: Proceedings of the 20th International Conference KES-2016.

[17] Derek L. Hansen, Ben Shneiderman, Marc A. Smith, and Itai Himelboim. 2020. Chapter 1 - Introduction to social media and social networks. In *Analyzing Social Media Networks with NodeXL (Second Edition)* (second edition ed.), Derek L. Hansen, Ben Shneiderman, Marc A. Smith, and Itai Himelboim (Eds.). Morgan Kaufmann, 3 – 10. https://doi.org/10.1016/B978-0-12-817756-3.00001-7

[18] Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE* 9, 6 (06 2014), 1–12. https://doi.org/10.1371/journal.pone.0098679

[19] Yuntao Jia, Victor Lu, Jared Hoberock, Michael Garland, and John C. Hart. 2012. Chapter 2 - Edge v. Node Parallelism for Graph Centrality Metrics. In

GPU Computing Gems Jade Edition, Wen mei W. Hwu (Ed.). Morgan Kaufmann, Boston, 15 – 28. https://doi.org/10.1016/B978-0-12-385963-1.00002-2

[20] Michael W. Kearney. 2019. *rtweet: Collecting Twitter Data*. https://cran.r-project.org/package=rtweet R package version 0.6.9.

[21] Vishal. A. Kharde and Sheetal. Sonawane. 2016. Sentiment Analysis of Twitter Data : A Survey of Techniques. *CoRR* abs/1601.06971 (2016). arXiv:1601.06971 http://arxiv.org/abs/1601.06971

[22] Efthymios Kouloumpis, Theresa Wilson, and Johanna Moore. 2011. Twitter Sentiment Analysis: The Good the Bad and the OMG! *ICWSM*.

[23] Shawn Martin, W. Michael Brown, Richard Klavans, and Kevin W. Boyack. 2011. OpenOrd: an open-source toolbox for large graph layout. In *Visualization and Data Analysis 2011*, Pak Chung Wong, Jinah Park, Ming C. Hao, Chaomei Chen, Katy Börner, David L. Kao, and Jonathan C. Roberts (Eds.), Vol. 7868. International Society for Optics and Photonics, SPIE, 45 – 55. https://doi.org/10.1117/12.871402

[24] Mark E. J. Newman and Michelle Girvan. 2004. Finding and evaluating community structure in networks. *Phys. Rev. E* 69 (Feb 2004), 026113. Issue 2. https://doi.org/10.1103/PhysRevE.69.026113

[25] Tahir M. Nisar and Man Yeung. 2018. Twitter as a tool for forecasting stock market movements: A short-window event study. *The Journal of Finance and Data Science* 4, 2 (2018), 101 – 119. https://doi.org/10.1016/j.jfds.2017.11.002

[26] Fabiola S. F. Pereira, Sandra de Amo, and João Gama. 2016. Evolving Centralities in Temporal Graphs: A Twitter Network Analysis. In *2016 17th IEEE International Conference on Mobile Data Management (MDM)*, Vol. 2. 43–48. https://doi.org/10.1109/MDM.2016.88

[27] Ioannis Pitas. 2016. *Graph-Based Social Media Analysis*. CRC Press. https://books.google.es/books?id=BvYYCwAAQBAJ

[28] John Price, Neil Farrington, and Lee Hall. 2013. Changing the game? The impact of Twitter on relationships between football clubs, supporters and the sports media. *Soccer & Society* 14, 4 (2013), 446–461. https://doi.org/10.1080/14660970.2013.810431 arXiv:https://doi.org/10.1080/14660970.2013.810431

[29] Hassan Saif, Yulan He, and Harith Alani. 2012. Semantic Sentiment Analysis of Twitter. In *The Semantic Web – ISWC 2012*, Philippe Cudré-Mauroux, Jeff Heflin, Evren Sirin, Tania Tudorache, Jérôme Euzenat, Manfred Hauswirth, Josiane Xavier Parreira, Jim Hendler, Guus Schreiber, Abraham Bernstein, and Eva Blomqvist (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 508–524.

[30] Wernard Schmit and Sander Wubben. 2015. Predicting Ratings for New Movie Releases from Twitter Content. 122–126. https://doi.org/10.18653/v1/W15-2917

[31] Shiladitya Sinha, Chris Dyer, Kevin Gimpel, and Noah Smith. 2013. Predicting the NFL using Twitter. *Proc. ECML/PKDD Workshop on Machine Learning and Data Mining for Sports Analytics* (10 2013).

[32] Michael Steketee, Atsushi Miyaoka, and Maura Spiegelman. 2015. Social Network Analysis. In *International Encyclopedia of the Social Behavioral Sciences (Second Edition)* (second edition ed.), James D. Wright (Ed.). Elsevier, Oxford, 461 – 467. https://doi.org/10.1016/B978-0-08-097086-8.10563-X

[33] Narges Tabari, Piyusha Biswas, Bhanu Praneeth, Armin Seyeditabari, Mirsad Hadzikadic, and Wlodek Zadrozny. 2018. Causality Analysis of Twitter Sentiments and Stock Market Returns. In *Proceedings of the First Workshop on Economics and Natural Language Processing*. Association for Computational Linguistics, Melbourne, Australia, 11–19. https://doi.org/10.18653/v1/W18-3102

[34] Marijn ten Thij, Tanneke Ouboter, Daniël Worm, Nelly Litvak, Hans Berg, and Sandjai Bhulai. 2015. Modelling of Trends in Twitter Using Retweet Graph Dynamics. https://doi.org/10.1007/978-3-319-13123-8_11

[35] Vincent A. Traag, Ludo Waltman, and Nees Jan van Eck. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Scientific Reports* 9, 5233 (2019).

[36] Jo Williams, Susan J Chinn, and James Suleiman. 2014. The value of Twitter for sports fans. *Journal of Direct, Data and Digital Marketing Practice* 16, 1 (01 Jul 2014), 36–50. https://doi.org/10.1057/dddmp.2014.36

[37] Yuto Yamaguchi, Tsubasa Takahashi, Toshiyuki Amagasa, and Hiroyuki Kitagawa. 2010. TURank: Twitter User Ranking Based on User-Tweet Graph Analysis. In *Web Information Systems Engineering – WISE 2010*, Lei Chen, Peter Triantafillou, and Torsten Suel (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 240–253.

[38] Reza Zafarani, Mohammad Ali Abbasi, and Huan Liu. 2014. *Social media mining: an introduction.* Cambridge University Press.