Tweets about Barcelona City Council Elections 2015

Master in Intelligent and Interactive Systems

A. Mestre[1] and A. Moral[2]

[1] anagabriela.mestre01@estudiant.upf.edu
[2] albert.lleo3@gmail.com

Abstract. The main goal of this paper is to analyze popularity on each political group leader against BeC - *Barcelona en comú*, winner of the elections, and include a focus on the segmentation and relation between BeC-p/Bec-m. To sum up, we want to perform a sentiment analysis on each cluster of all political parties including division of BeC to study the content of each tweet that a political party leader made in that concrete situation.
This study is based on the *When a Movement Becomes a Party: Computational Assessment of New Forms of Political Organization in Social Media* paper. Their work examined whether Barcelona en Comú preserved a decentralized structure or adopted a conventional centralized organization

Keywords: elections · popularity · influence · segmentation · 15 M.

1 Objectives

The main goal of this project is to perform an study on the *"Tweets about Barcelona City Council Elections 2015"* dataset and subtract some features to obtain the aimed objectives. The mentioned main objectives are the ones lists below:

- Analyze the popularity of every political organization leader.
- Perform a comparison between the elected political party leader *@adacolau* (BeC, Barcelona en comú) and the rest of leaders.
- Based on the *When a Movement Becomes a Party* Paper. They found that in Barcelona en Comú two well-defined groups co-existed: a cluster dominated by the party leader and the collective accounts (BeC-party, BeC-p), and another cluster formed by the movement activists (BeC-movement, BeC-m). Therefore, we want to try to localize and compare these two main clusters inside of the BeC political party.
- Perform a sentiment analysis on the text of all tweets involved. This follows a purpose of study the content of each tweet that a political party leader

made in that “crispy” situation.

2 Data

The dataset *”Tweets about Barcelona City Council Elections 2015”* consists of a collection of tweets in relation to the campaign for the 2015 Barcelona City Council election (May 1-26, 2015). The main political organizations involved on the campaign were the following:

- Barcelona en Comu (BeC)
- Convergencia i Unió (CiU)
- Ciudadanos (Cs)
- Capgirem Barcelona (CUP)
- Esquerra Republicana de Catalunya (ERC)
- Partit Popular de Catalunya (PP)
- Partit dels Socialistes de Catalunya (PSC)

The fields that can be subtracted of this dataset are the ID of the tweet, the text of each tweet, the hashtags, ID of the original tweet if that tweet is a retweet, media URLs, entity URLs and timestamp in ms.

3 Methodology

In order to succeed on all the objectives listed above we have to first perform the data extraction from the dataset in order to achieve our goals. This task will be accomplished by creating a python script that generates a co-mention network. This means that we are building a weighted graph where the nodes are the usernames that appear on all tweets from the dataset, and every edge between two nodes imply that those two usernames appear on the same tweet. The weight for every edge represents the number of times two usernames appear on a tweet.

Secondly, visualize our data by applying clustering to all the tweets in order to split our data into all the political groups involved in the elections, and then apply different centrality metrics in order to observe the most relevant component in each cluster we found.

Once we have subtracted all the information, we can proceed to perform the comparison between the political leaders. In addition, we will also try to localize BeC-p and BeC-m in our splitted data and compare the popularity between them.

To conclude, we will use the Spanish sentiment analysis library in python *senti-py* on each political party to obtain the social opinion on the elections of Barcelona council 2015 by subtracting the text of each tweet published.

4 Results

Once we substracted our co-mention network from the provided dataset and uploaded it to gephi, we could see our resultant graph as the following figure:

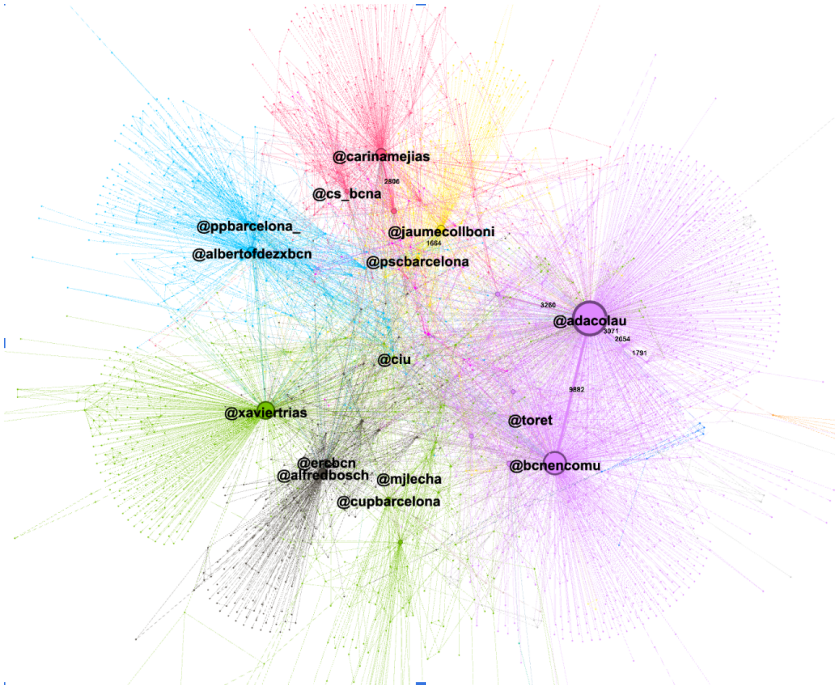

Fig. 1. Resultant graph by extracting the data features

Once we analyzed our graph, we could substract the following results:

- **Clustering performed**

  After creating the co-mention network, we could apply modularity to our graph (using a standard value of resolution 1) to obtain 19 different communities. For this different communities, we clearly could identify the main 6 refearing to the political organisations. The results were not exactly as expected, due to we had an error by differentiating the political parties of

@ciu and @cupbarcelona. This could have happened because these two political parties are a bit related between one another. They both share similar objectives and therefore, our twitter-generated network classified them into the same cluster.
Regarding all other parties, we were able to clearly identify all the different clusters related to each political party.
We reaffirm the correct clustering results by adding labels on the nodes with the name of the political candidates of each political party, this procedure helped us see that each node with that label represented the nucleus of a cluster.

- **Popularity**

  Once we substracted the different clusters, we tried different metrics as Pagerank, modularity class, Eccentrycity or Betweness Centrality but finally, we applied a partition of the node size by its degree. With this method we could identify the nodes with the highest degree, therefore the most popular nodes.
  As expected, we get that the most popular node is *@AdaColau* with her political organisation *@bcnencomu* followed by the former mayor *@xaviertrias* and his organisation *@ciu* and on the third place we could find *@carinamejias* with her political org. *@Cs*.

- **Analysis on BeC-p and BeC-m**

  Regarding the identification, analysis and differentiation of the two movements within the political group @BeC (Barcelona en Comú), which are BeC-m and BeC-p already mentioned above, we have not been able to make a clear identification of both due to that when calculating the modularity, we include both groups within the same cluster.
  To try to differentiate them we have carried out multiple calculations of the modularity changing the resolution to obtain a different number of communities and thus differentiate them, but without success. We even tried another method, applying Girvan-Newman clustering to our graph but also with no different results.

- **Sentiment Analysis**

  We subtracted the text field of each tweet made by all the candidates to political leader and we performed a sentiment analysis on it to try to identify the tendency on their writing at that moment.
  Find below the corresponding results per every political party. As we can see, most of them have an overall negative score, like Ada Colau (Barcelona en comú) with a 51.61% of negative tweets, Xavier Trias (CiU) with a 57.31%, Carina Mejías (Ciutadans) with a 50.26%, Alberto Fernandez (PP) with a 51.07% and Jaume Collboni (PSC) with a negative score of 46.06%. These

high negative scores could be related with the fact that as political parties they tend to criticize multiple actions. Another fact for these numbers could be the Catalan tweets. The sentiment analysis library is intended to be used with Spanish text and some of our tweets are written in Catalan. However, since Spanish and Catalan are quite similar, we have observed that the performance for both cases are very much alike. Nevertheless, since *sentipy* is not a Catalan tool, some mistakes could be noticeable.

Table 1. Sentiment Analysis on every leader tweet

Candidate	Positive	Negative	Neutral
@adacolau [BeC]	30,65%	**51,61**%	17,74%
@alfredbosch [ERC]	**57,38**%	24,76%	17,85%
@carinamejias [Cs]	35,36%	**50,26**%	14,47%
@mjlecha [CUP]	**50,0**%	38,8%	11,2%
@xaviertrias [CiU]	29,87%	**57,31**%	12,82%
@albertofdezxbcn [PP]	39,79%	**51,07**%	16,13%
@jaumecollboni [PSC]	34,53%	**46,04**%	19,41%

From this table we can subtract that the top three most popular candidates have a higher negative percentage in their published tweets. We have no way to justify this fact, the only possible way would have been to get the tweets of the rest of the population, and not the tweets of the candidates to truly see their opinion about them and therefore, relate them to the popularity obtained before.

5 Conclusions

To conclude our project, we state that Ada Colau, candidate of Barcelona en Comú political organisation (BeC), is the most popular leader of all the possible candidates, followed by Xavier Trias (CiU) who is the former mayor of Barcelona and Carina Mejias, candidate of Ciudadanos (C's) organisation.

Now, focusing on the identification and segmentation of the two main groups inside BeC, we were not able to separate clearly these two main clusters by applying only modularity.

Finally, in a general way we can state that the tweets published by the political leaders follow a negative tendency, maybe due to the tension related to

the elections, and their political campaign could be criticizing the other political organisations and candidates.