Political Polarization in Participatory Media

David Garcia, Frank Schweitzer

May 4, 2013

Online participatory media, such as social networking sites or discussion forums, play a key role in the current political climate, offering the chance to gather large datasets of voter expression and behavior. A common feature of traditional mass media and online participatory media is that both offer politicians a way to reach a very large numbers of users in very short time. But the main difference is that in the latter, users can also reach large numbers of other users by posting publicly accessible posts and comments. This leads to collective patterns that are emergent phenomena resulting from the interaction of many users. One of these is the polarization of emotions and opinions which is difficult to explain by analyzing the "representative user", due to the inherent heterogeneity of citizens in terms of political topics. Despite these individual differences, the networks of `Twitter` users [1], and the dynamics of video views in `Youtube` [3], and votes in `Reddit` [10] show the existence of statistical regularities of polarization.

We illustrate our approach analyzing a dataset from `Youtube`, composed of the view statistics and comments for the videos of the U.S. presidential campaigns of 2008 and 2012. Using sentiment analysis, we quantify the collective emotions expressed by the viewers of the videos of each campaign. Second, we present our study of cultural dynamics through the votes in the Eurovision song contest. We define a measure of polarization in the voting biases of the participants, and show its relations to economical indicators of Europe.

Under which conditions positive and negative emotions in online discussions amplify polarization in online participatory media? What is the role of user interaction in the formation of opinions, and how are these reflected in the political climate? How do economical and political factors influence the cohesion of societies as a whole? We address these questions in a quantitative way by means of novel computational methods like the automated emotional classification of texts, machine learning, statistical analysis of large-scale data, and agent-based modeling.

1 Polarization in Youtube campaigns

We analyzed the `Youtube` channels of the candidates for the US presidential elections in 2008 and 2012 [3], retrieving information for all videos in the official channels for Barack Obama (`barackobamadotcom`), Mitt Romney (`mittromney`), and John McCain (`johnmccaindotcom`). In total, 3753 videos where available at the date of the data retrieval, November 4th 2012. For each video, we gathered the available set of comments, composing a dataset of 474536 comment texts.

We divide our dataset in four subsets referring to both candidates in the campaigns of 2008 and 2012.

We processed the text of each comment using SentiStrength [9], a sentiment analysis tool for the extraction of emotional content from short texts. SentiStrength is the state-of-the-art tool for lexicon-based analysis of short, informal text. This is the case of `Youtube` comments, for which SentiStrength's accuracy is above 88% when classifying positive and negative text [9]. We used SentiStrength's output [3] to classify each comment c as positive ($e_c = +1$), negative ($e_c = -1$), or neutral ($e_c = 0$) .

We aggregate the emotions expressed in the set of comments of a video C_v by calculating the ratios of positive $P_v = \frac{\sum_{c \in C_v} e_c = 1}{|C_v|}$, negative $N_v = \frac{\sum_{c \in C_v} e_c = -1}{|C_v|}$, and neutral $U_v = \frac{\sum_{c \in C_v} e_c = 0}{|C_v|}$ comments. These three values allow us to map each video in a space that can equally represent videos that did not elicit collective emotional responses (high U_v), created positive or negative collective emotions (high P_v or high N_v), or created polarized emotional reactions (both high P_v and N_v). In Fig. 1 we plot each video as a point inside a triangle with a distance to the vertices inversely proportional to U_v for the upper vertex, N_v for the lower left vertex, and P_v for the lower right vertex. The collective emotional reaction to a video is classified as neutral (black), positive (green), negative (red), or undetermined (yellow), through a set of statistical tests [3].

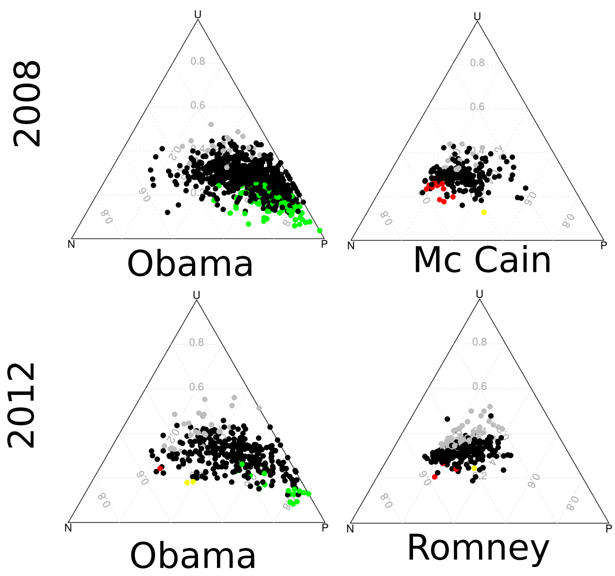

Figure 1: Triangular representation of the collective emotions in the comments for the videos in the four studied campaigns [3].

Focusing on the 2008 campaigns, the emotions expressed in the comments of Barack Obama's videos were clearly skewed towards positive expression, having numerous instances of positive collective responses and no case of a negative collective emotion. On the other hand, John Mc Cain's videos were triggering much more negative discussions in the few videos uploaded in his channel, having some cases of negative collective emotions. Both Obama's campaigns show a similar pattern towards positive emotions in 2008 and 2012, but they also include polarized collective reactions closer to the lower edge of the triangle. In the case of Mitt Romney's channel, we did not find any case of positive collective emotions, and the distribution of emotional expression seems similar to John Mc Cain's without reaching very high values of negativity. This kind of quantification of the type of emotional response of the `Youtube` videos of a political campaign opens the question of whether it is always desirable for a politician to attract large attention to these videos. Popularity does not need to be positive popularity, and online campaigns can indeed encourage negative discussions about the candidate.

2 Cultural polarization and debt crisis

Measuring polarization at an international level poses a challenge, as linguistic and cultural barriers complicate the gathering and analysis of datasets for such purpose. A notable example of international cultural expression is the Eurovision song contest, a song competition across European countries. The votes of this contest are decided by a combination of jury decisions and televoting, which allows all the viewers to choose their favourite song. The statistical patterns of this data have been shown to relate to cultural differences [6] and socioeconmical phenomena [5]. Here, we present how we measure cultural polarization through the contest results, and a statistical relation to economical indicators of the Euro crisis.

First, we define a measure of cultural affinity, which can be estimated from the raw Eurovision scores of a contest. We define the Friend-or-Foe (FoF) coefficient of country c_v towards country c_c, as estimated from a particular edition of the contest: $Fof(c_v, c_c) = \frac{p_{v,c}}{12} - \frac{s_c - p_{v,c}}{12(N-2)}$ where $p_{v,c}$ are the points assigned to c_c by c_v, s_c is the final score of c_c, and N is the total amount of countries voting in the studied edition of Eurovision. The Friend-or-Foe coefficient measures the overvoting or undervoting bias from a country to another, correcting for "song quality" as estimated by the final contest result. This way, we aim at removing the effects of the artistic component of the contest, highlighting the political or cultural biases that are commonly claimed to exist in Eurovision. If a country c_v assigns 12 points to c_c, while all the others assign 0, then $FoF(c_v, c_c) = 1$, which would be the maximum value of an overvoting bias. If c_v assigns 0 points to c_c but all the other countries assign 12, then $FoF(c_v, c_c) = -1$, representing the maximally negative Friend-or-Foe coefficient given the contest rules.

To measure cultural polarization, or the overall level of "disagreement" across participants, we calculate $Pol(t) = \sqrt{\frac{1}{E_t} \sum_{c,v} (FoF_t(c_c, c_v) - \langle FoF_t \rangle)^2}$, which is essentially the standard deviation of FoF across all the pairs of countries in the network, which amount to E_t. This polarization metric takes higher values when voting biases are strong, in comparison to the artistic component of the contest. If all countries agreed on the best songs in the same manner, the polarization would have a value close to zero.

In our analysis, we focus on a subset of countries called the EU-15, which is the set of members of the European Union since 1995. A careful observation of the time series of polarization in the

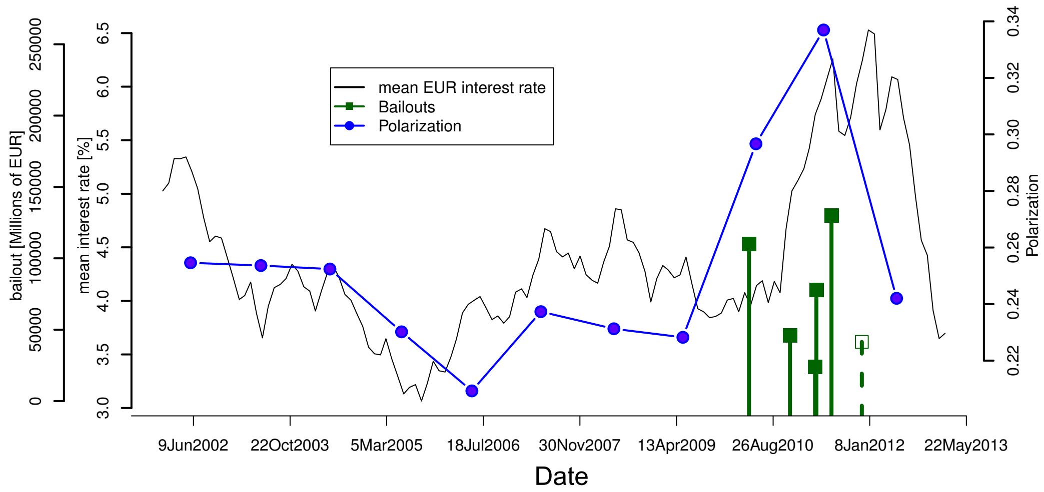

Figure 2: Time series of EU-15 polarization (blue dots) and monthly mean sovereign bond interest rate for Euro-funding countries (black line). Green squares indicates dates and volumes of bailout loans to EU countries (Greece, Portugal, Ireland) until January 2013. The last green square indicates the bailout to the Spanish private banking sector. (Figure from [5])

EU-15 in Figure 2 reveals a peaked value in 2010 and 2011, coinciding with the loans and austerity measures in Portugal, Ireland, Italy, Greece, and Spain. As a comparison, the polarization keeps relatively stable between 1997 and 2009. We relate this measure of polarization with the European debt crisis through the interest rate of the sovereign bonds of the countries of the Eurozone. Figure 2 shows the mean interest rate of the long-term sovereign bonds of the 12 Euro founder countries, all part of the EU-15. Both polarization and interest rate jointly increase in 2010 and 2011, seemingly increasing polarization before the interest rate. This analysis is supported by statistical tests [5], finding maximal correlation (0.894) between Polarization and mean interest rate 7 months later.

The relation between these two variables can be interpreted as support for the theory that both are influenced by a third component, or that both are manifestations of the same phenomenon. This way, the polarization in Eurovision would be an early indicator for a social and cultural phenomenon, which is followed by states of distrust in the economy of the involved countries. Our polarization metric provides a way to quantify how *society* reacts to political decisions and the crisis in general, in a similar manner as sovereign bond interest rates measure how *the market* reacts to the same phenomenon.

3 Future research

One step further, we also aim at reproducing these patterns by agent-based models based on emotional influence [7]. Agent-based models are starting to be used as a tool in social psychology and emotion research [8]. We have applied these kind of models to reproduce certain patterns of emotional interaction in chatrooms [2], and product reviews [4]. Agent-based models provide a tractable link between the collective human behavior and the interactions of individuals, which is of particular value for political sciences.

Agent-based models are based on assumptions that describe the individual behavior and the interaction between agents. These assumptions need to be empirically testable in order to be integrated in a larger scientific perspective, and to be applied in real-world applications. Having access to the online traces of users in social networks is usually not enough to design the model at a fine-grained level. In our approach we go beyond the publicly visible layer of the user, integrating other data sources of individual behavior. With respect to political applications, we plan to test our models with data on individual activity in voting advice applications (VAAs). VAAs match the policy preferences of real voters (who visit the website) with that of candidates and/or political parties (that are already encoded in the online system), through the voter's answers to a set of predefined questions. While designed to provide advice to the users, VAAs also provide an experimental platform to gather data on voter preferences, which cannot be easily retrieved through online traces or election surveys.

References

[1] M. D. Conover, B. Gonc, A. Flammini, and F. Menczer. Partisan Asymmetries in Online Political Activity. *EPJ DataScience*, 1(6):1–17, 2012.

[2] A. Garas, D. Garcia, M. Skowron, and F. Schweitzer. Emotional persistence in online chatting communities. *Scientific Reports*, 2:402, May 2012.

[3] D. Garcia, F. Mendez, U. Serdult, and F. Schweitzer. Political polarization and popularity in online participatory media : An integrated approach. In *Proceedings of the first edition workshop on Politics, elections and data - PLEAD '12*, pages 3–10, 2012.

[4] D. Garcia and F. Schweitzer. Emotions in Product Reviews–Empirics and Models. In *SocialCom'11*, pages 483–488, 2011.

[5] D. Garcia and D. Tanase. Measuring cultural dynamics through the eurovision song contest. *ACS - Advances in Complex Systems (Submitted)*, 2013.

[6] V. Ginsburgh and A. G. Noury. The Eurovision Song Contest. Is voting political or cultural? *European Journal of Political Economy*, 24(1):41–52, Mar. 2008.

[7] F. Schweitzer and D. Garcia. An agent-based model of collective emotions in online communities. *The European Physical Journal B*, 77(4):533–545, 2010.

[8] E. R. Smith and F. R. Conrey. Agent-based modeling: a new approach for theory building in social psychology. *Personality and social psychology review : an official journal of the Society for Personality and Social Psychology, Inc*, 11(1):87–104, Feb. 2007.

[9] M. Thelwall, K. Buckley, and G. Paltoglou. Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, 63(1):163–173, Jan. 2012.

[10] P. Van Mieghem. Human Psychology of Common Appraisal: The Reddit Score. *IEEE Transactions on Multimedia*, 13(6):1404–1406, Dec. 2011.