Tlatelolco LAB: Identification of communities and narratives during focal events on Twitter

Martín Zumaya [1] [2]   Diego Espitia [1]   Luis Angel Escobar [1]

[1]Programa Universitario de estudios sobre Democracia, Justicia y Sociedad, UNAM, Mexico
[2]Centro de Ciencias de la Complejidad, UNAM, Mexico

Abstract

Tlatelolco-Lab is an **interdisciplinary** team based at UNAM, dedicated to the investigation and analysis of short and long-term events around key public debates in the political and social context of contemporary Mexico.

Thanks to our interdisciplinary approach, we show how statistical methods can be used alongside network science and sociological methodologies, to depict a landscape of the political dynamics and traffic manipulation mechanisms in online social networks.

In this poster we briefly describe our **general methodology** to study the activity on Twitter and identify **the main narratives around relevant political and social events**.

Actions and interactions between accounts

Users on Twitter can perform a set of different actions on other users updates, each one of them can be thought as interactions between accounts with different purposes.

- **Retweet:** This action **amplifies or spreads** the content of the original publication.
- **Quote:** This action is similar to the retweet interaction, in the sense of amplification or spreading, but **includes a comment to the original publication.**
- **Reply:** This action **replies or reacts to a publication**, but doesn't necessarily imply an amplification or spread of the original content.
- **Like:** This action expresses **endorsement** of a user to the publication or content posted by other users. It doesn't necessarily implies amplification or spread of the original content.

This set of actions that users can perform, allow us to define **interaction networks** between them, so that **nodes** correspond to users or accounts and **links** to one of the possible interactions.

Focal events and narratives

We focus on the **identification of the main narratives during focal events** in social media, specifically on Twitter.

- **Focal events:** High activity moments in social media around relevant political or social events occurring "offline".
- **Narratives:** Narratives correspond to the different **interpretations and meanings** that sets of users give to the content and the debate in online social networks around focal events.

We can classify narratives around focal events in two categories:

- **Main Narratives:** These are the narratives promoted by **opinion leaders or relevant accounts** that represent specific views or readings of focal events. Main narratives are generated mainly by **tweets or original publications** of "influencers" or relevant accounts in the platform. These narratives are **amplified** mainly by **retweets and quotes**.
- **Response Narratives:** Correspond to the reaction of sets of users to the main narratives in a focal event. These narratives usually represent antagonistic views to main narratives and are driven mainly by **quotes and replies**.

Communities and narratives characterization

Narratives in online social networks can be identified by two main aspects:

- **Communities in interaction networks between users:** Interactions in online social networks organize around **content and affinities**. Components or communities in e.g. retweet interaction networks represent **sets of likely-minded users around views of focal events**.
- **Content:** The content produced and amplified in communities or components in interaction networks represent the **interpretation and ideological views of users about the matters discussed during focal events**.

General methodology description

The **main methodological steps** for both the communities and narratives identification are the following:

- **Data collection:** Once a focal event is identified, we collect data from the **Twitter API**, either by filtering a real-time stream of data with **specific keywords or hashtags**, or by a search in the Twitter's full archive around a **period of time and relevant keywords and hashtags**.
- **Interaction network construction:** We classify the collected data around the focal event with respect with the kind of interaction they represent (retweet, reply, ...), and build the weighted interaction network between users.
- **Community detection:** We identify the communities in the interaction network via community detection algorithms, such as the Girvan-Newman or Louivain algorithms.
- **Content characterization:** We filter the data of each community and treat it separately, and analyze it with: **topic modelling using Latent Dirichlet Allocation (LDA), hashtags co-occurrence networks** and **shared media such as photos or images**.
- **Data interpretation:** Once communities and their content are identified, the results are analyzed to interpret and identify the present narratives around the focal event.

Communities in retweets networks

Communities identified by the Girvan-Newman algorithm in retweet networks around the midterm elections in Mexico, correspond to groups of political and news media around the main mexican political parties.
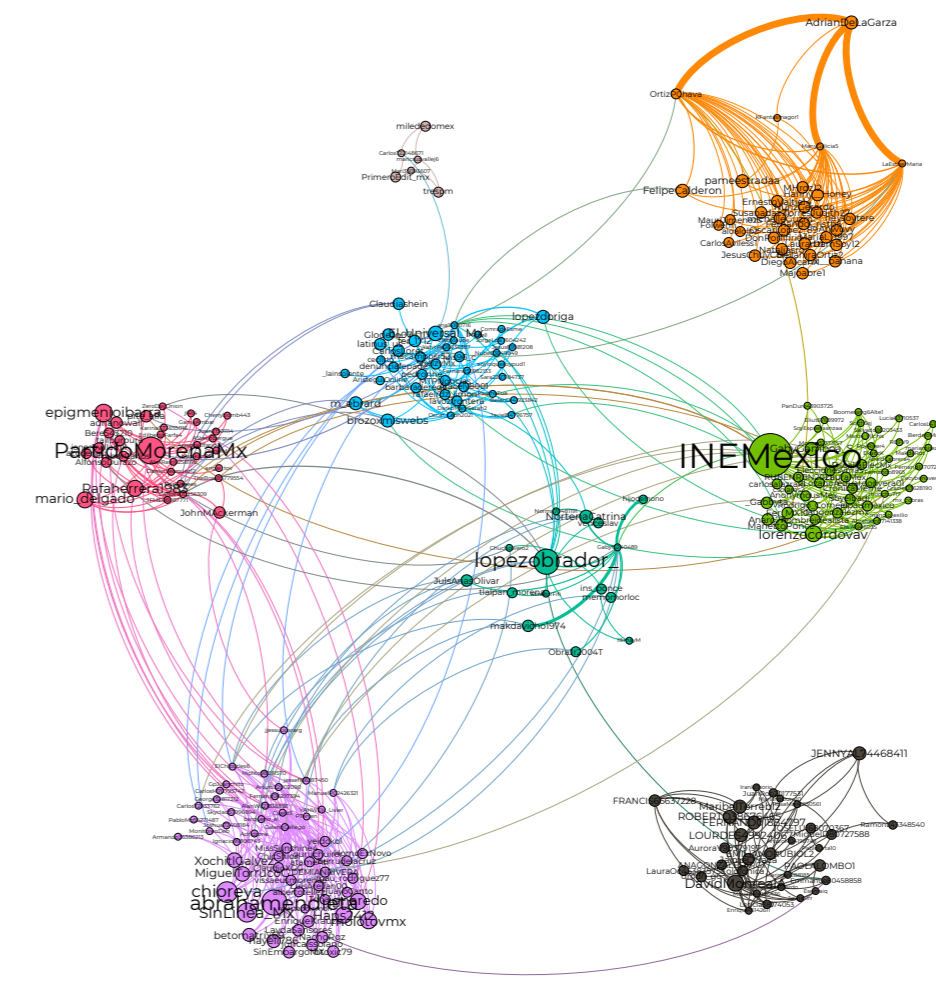

Figure 1. Identified communities during the midterm election day in June 2021 in Mexico. Communities map to the main mexican political parties.

Hashtag co-occurrence networks

The communities in the co-occurrence network of the hashtags used by right wing against abortion, reflect how they reacted to the abortion decriminalization in Mexico City and the different groups that took place in the conversation, for example, physicians and medical workers, and religious oriented accounts. The criticism against magistrates had an important presence in the debate.
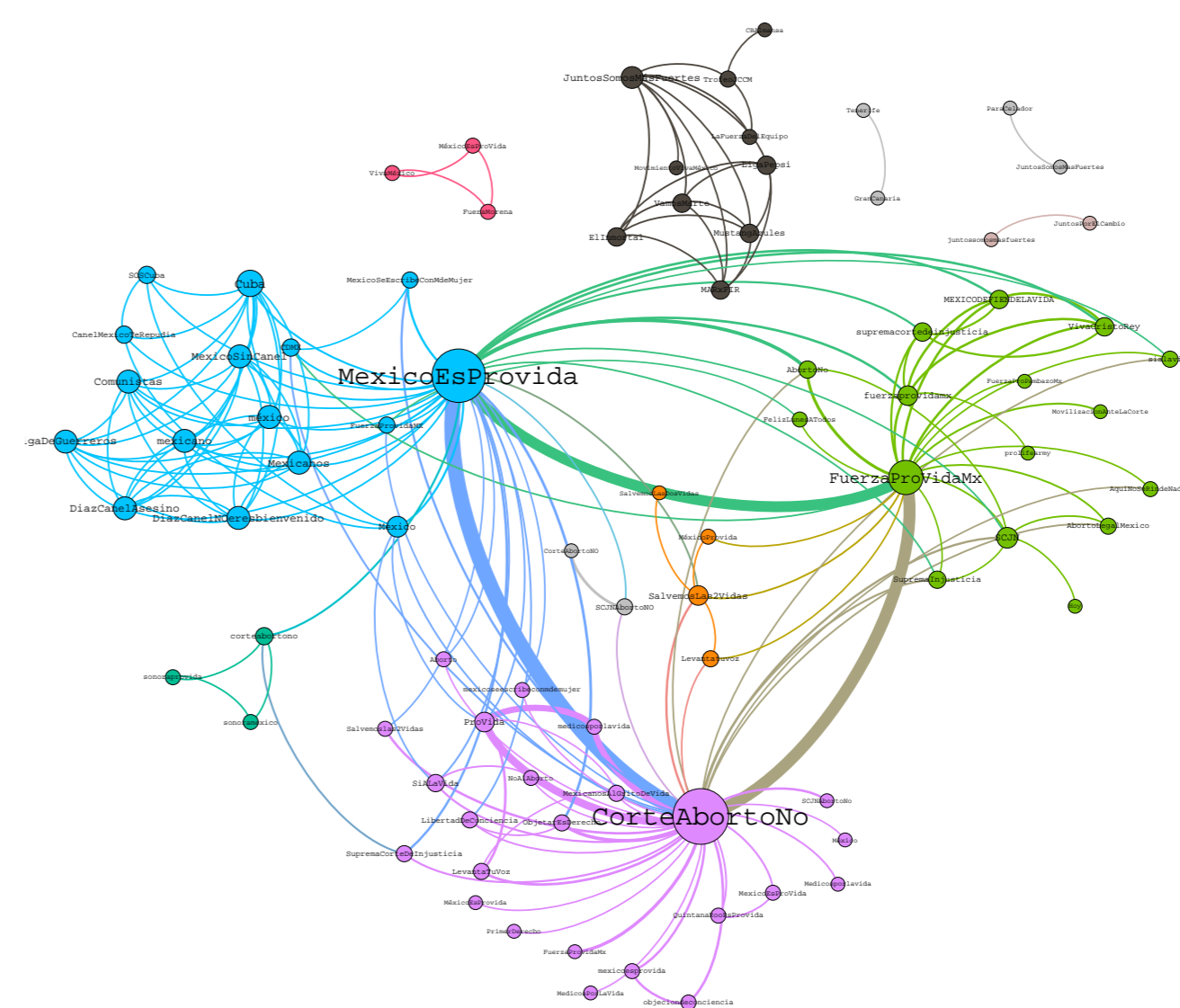

Figure 2. Co-occurrence network of hashtags used by prolife and right wing Twitter accounts in the context of the abortion decriminalization in Mexico City. Communities map to arguments against abortion.

Topic modelling and social movements criminalization

Topic modelling of tweets' texts using LDA in the context of demonstrations of elementary school teachers in Mexico City, reveals a set of twitter users whose main narrative is the criminalization of protesters and demonstrations.

Figure 3. Topic modelling of tweets' texts using LDA in the context of demonstrations of elementary school teachers in Mexico City. The criminalization perspective of the texts is evident by the use of words insulting and comparing the protesters with criminals (*vándalos* and *delincuentes* in spanish).