Tracking the political opinion landscape in Twitter during electoral periods

T. Mussi Reyero, M.G. Beiró, J.I. Alvarez-Hamelin, L. Hernández and D. Kotzinos

CONICET

.UBA fiuba FACULTAD DE INGENIERÍA

LPTM Laboratoire de Physique Théorique et Modélisation

etis Equipes Traitement de l'Information et Systèmes

I

INTRODUCTION

Aim

- The scope of this work is to design a mechanism to track election processes in Twitter.
- We applied it to the Argentine presidential elections of 2015 and 2019.
- We were interested in:
  - automatically detecting topics of discussion in the social network.
  - characterizing supporters based on the topics they talk about.
  - measuring similarities across different parties and cohesion inside a party.
  - determining if the evolution of these similarities can reveal changes in the political scenario.

Related work

- Usual works concerning elections in Twitter usually pick specific hashtags or keywords in order to extract tweets.

- Or either they track the accounts of specific public figures.

- Here we propose a different approach: to capture all the information from the candidates' followers to compute statistics of their behavior.

Data capturing

Before each election, we check the Twitter accounts from the main candidates or parties. We track who their followers are, and we capture all the followers' tweets during the election process.

By restricting ourselves to active Argentinian followers, we obtained the following amount of data:

Election	Number of tracked users	Number of processed tweets
2015	218k	50M
2019	586k	280M

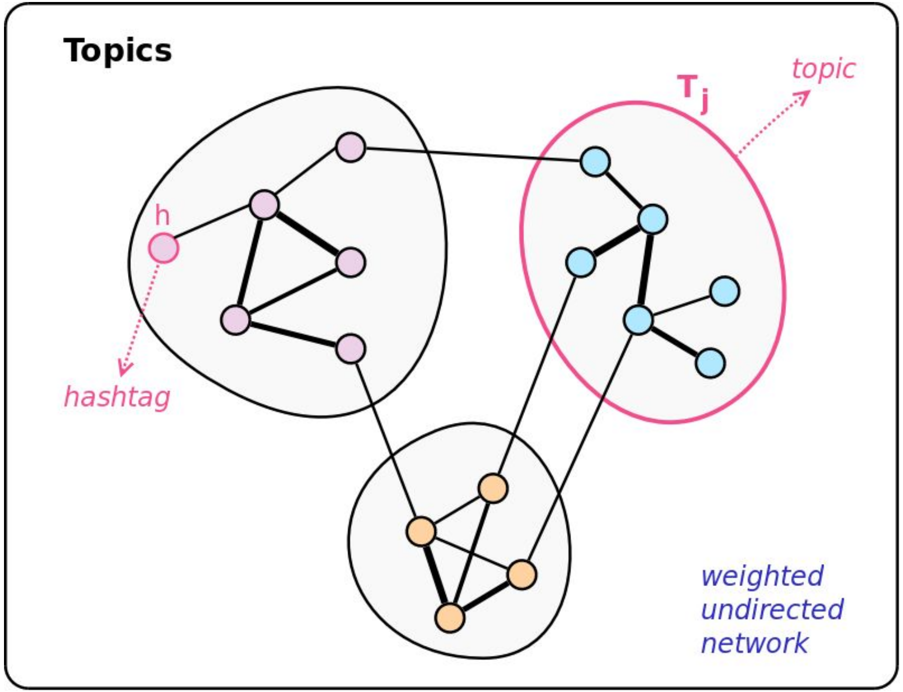
DEFINING

TOPICS

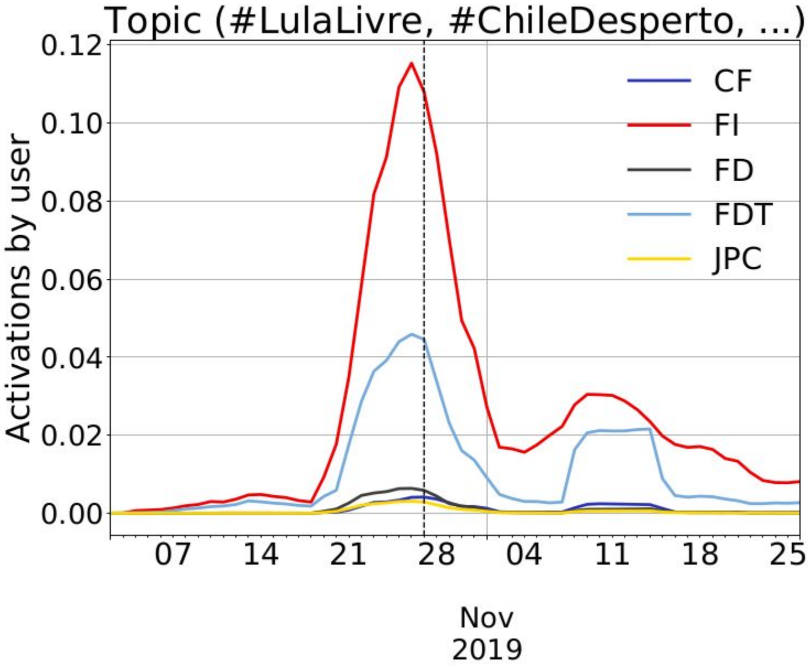
II

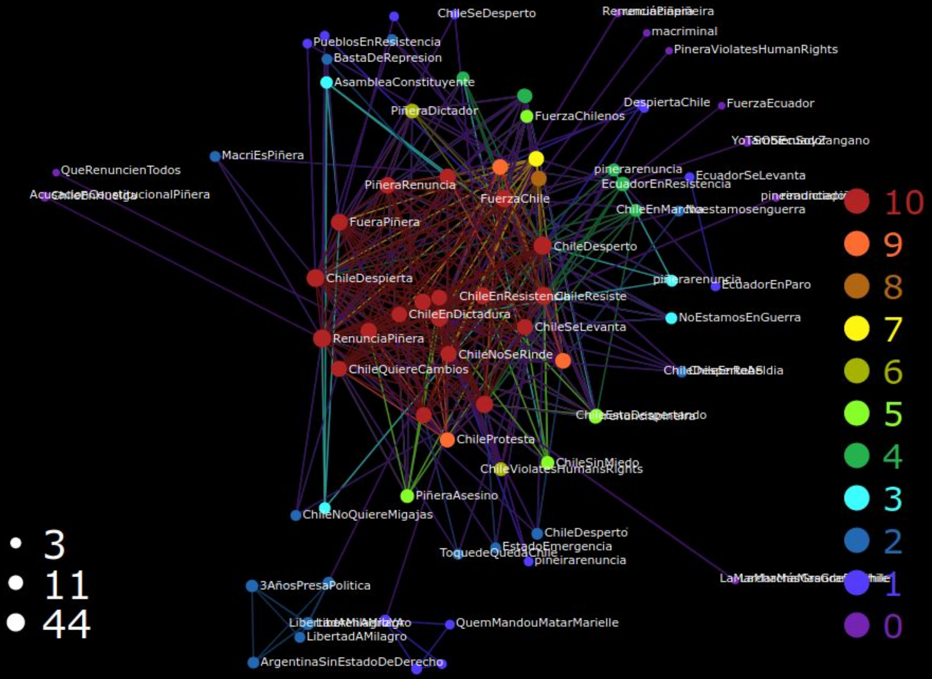
Topic discovery

- In order to detect topics, we build a weighted undirected network of hashtag co-occurrence.
- Then, we apply a community detection algorithm to this network.



Example


Topic (#LulaLivre, #ChileDesperto, ...)
Activations by user
0.12
0.10
0.08
0.06
0.04
0.02
0.00
CF
FI
FD
FDT
JPC
07
14
21
28
04
11
18
25
Nov
2019

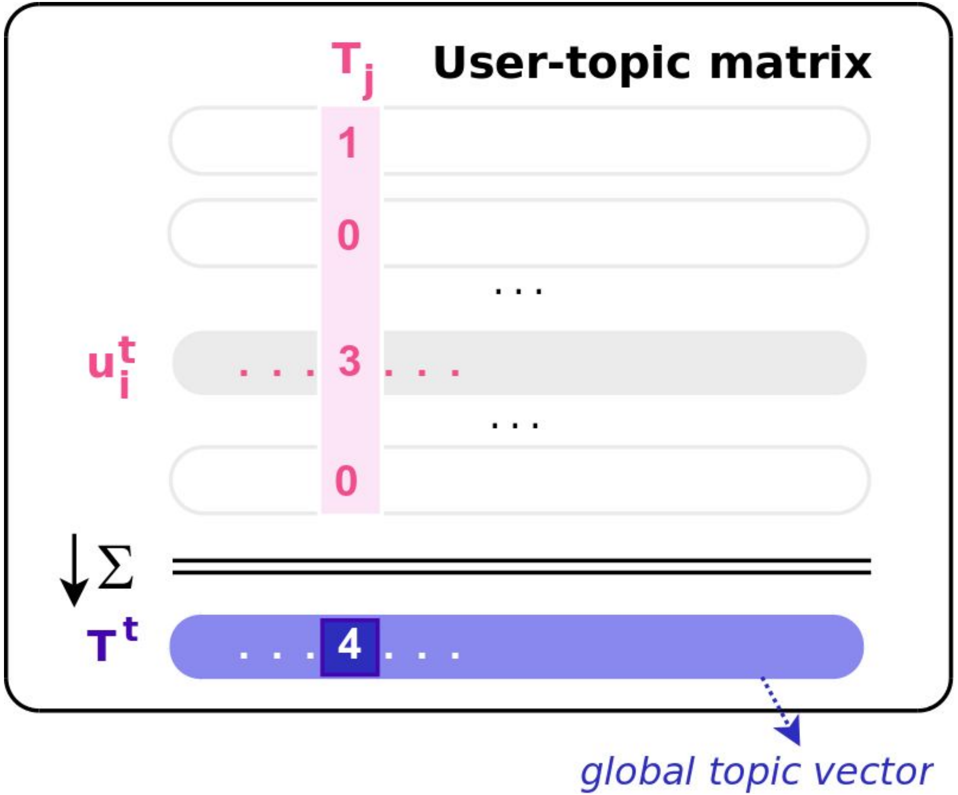

ChileSeDesperto
macriminal
PineraViolatesHumanRights
PueblosEnResistencia
BastaDeRepresion
AsambleaConstituyente
PiñeraDictador
FuerzaChilenos
DespiertaChile
FuerzaEcuador
MacriEsPiñera
QueRenuncienTodos
pinerarenuncia
EcuadorSeLevanta
EcuadorEnResistencia
PiñeraRenuncia
FuerzaChile
FueraPiñera
ChileDesperto
ChileDespierta
EcuadorEnParo
ChileEnResistencia
ChileResiste
ChileEnDictadura
NoEstamosEnGuerra
ChileSeLevanta
RenunciaPiñera
ChileNoSeRinde
ChileQuiereCambios
ChileProtesta
ChileSinMiedo
PiñeraAsesino
ChileNoQuiereMigajas
ChileDesperto
EstadoEmergencia
pineirarenuncia
3AñosPresaPolitica
LibertadAMilagro
QuemMandouMatarMarielle
ArgentinaSinEstadoDeDerecho
3
11
44
10
9
8
7
6
5
4
3
2
1
0

III

CHARACTERIZING USERS

User description vectors (Step 1/3)

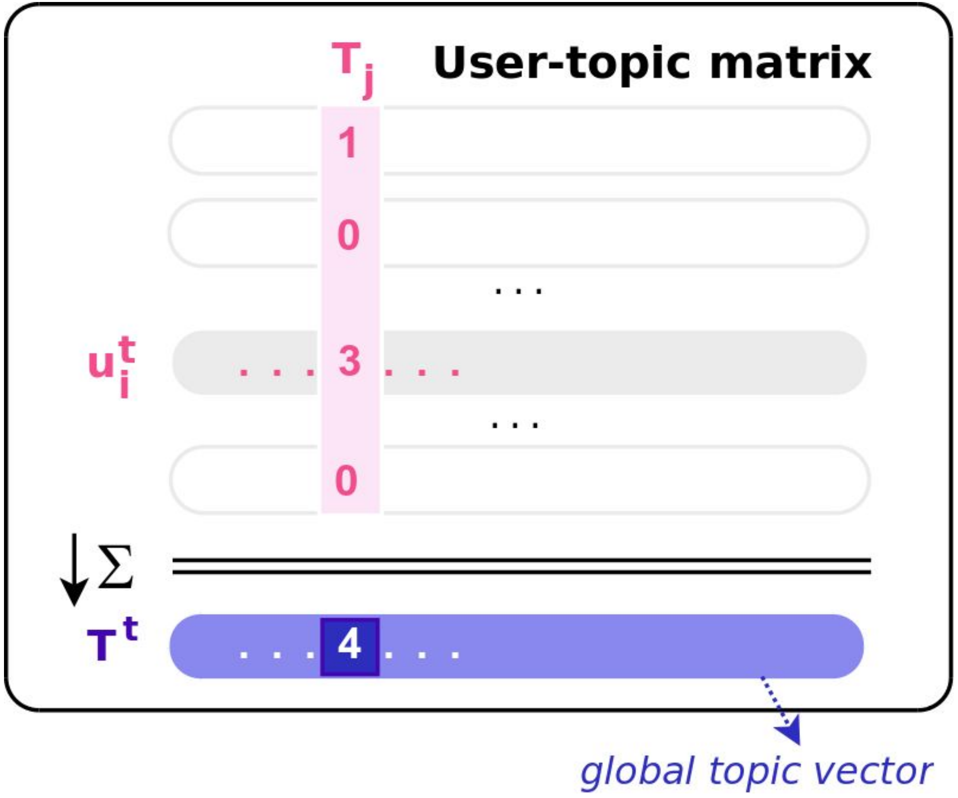
- In a time window of 7 days, we measure the **activity** of the users across the discovered topics.
- This activity is measured in terms of the number of times that a hashtag contained in the topic is used in a tweet, or is contained in a retweet.
- Thus, each user is characterized by a vector of size equal to the number of topics.



User description vectors (Step 2/3)

- We use the global topic vector to normalize the user vector.

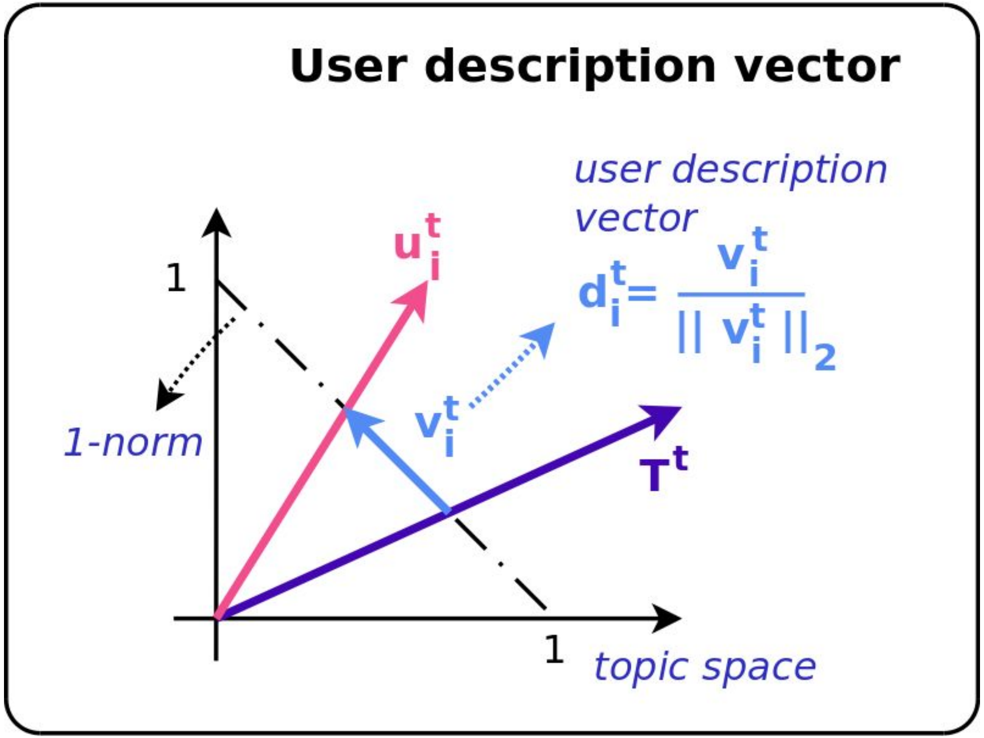
- But instead of using a technique like TF-IDF, we compute the difference between the user activity vector (rescaled) and the global topic vector (rescaled).



User description vectors (Step 3/3)

- This result is also rescaled using the L2-norm.

- The final **user description vector** gives an idea of the deviation of the user's interests with respect to the general interests of the Twitter community.



MEASURING SIMILARITIES ACROSS GROUPS

IV

Intra-group and inter-group similarity

- The normalized description vectors previously defined allow us to measure the similarity between two users' interests as the cosine of the angle between their description vectors.

- This same measure allows for defining cohesion of a **group** of users (or the dissention across **different groups)**, as the average similarity among all pairs of members.

- A little algebra reveals that these group measures can be easily computed as the cosine similarity between representative agents (whose description vector is the average of the description vectors of all the group users).

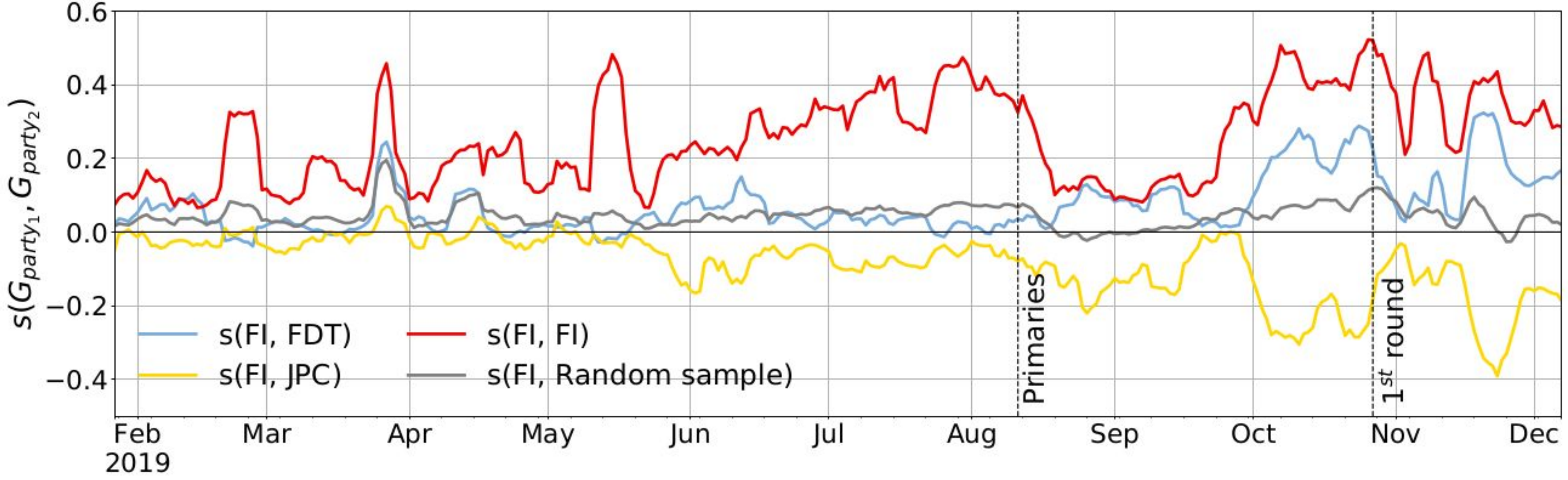
These definitions allowed us to observe different aspects of the electoral processes.

V

RESULTS

Result #1: Similarities reveal closeness between political parties

2019 elections



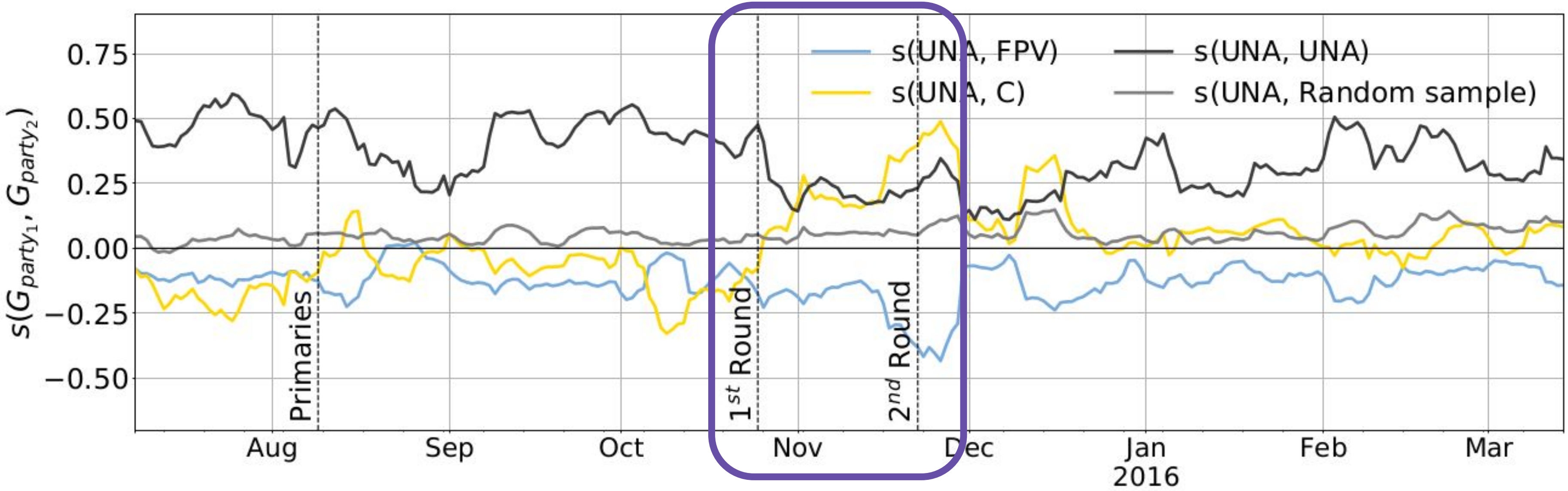
Left party

Center-left party (peronism)

Right party

Result #2: Similarities can also reveal the reorientation of some supporters after a political event

2015 elections

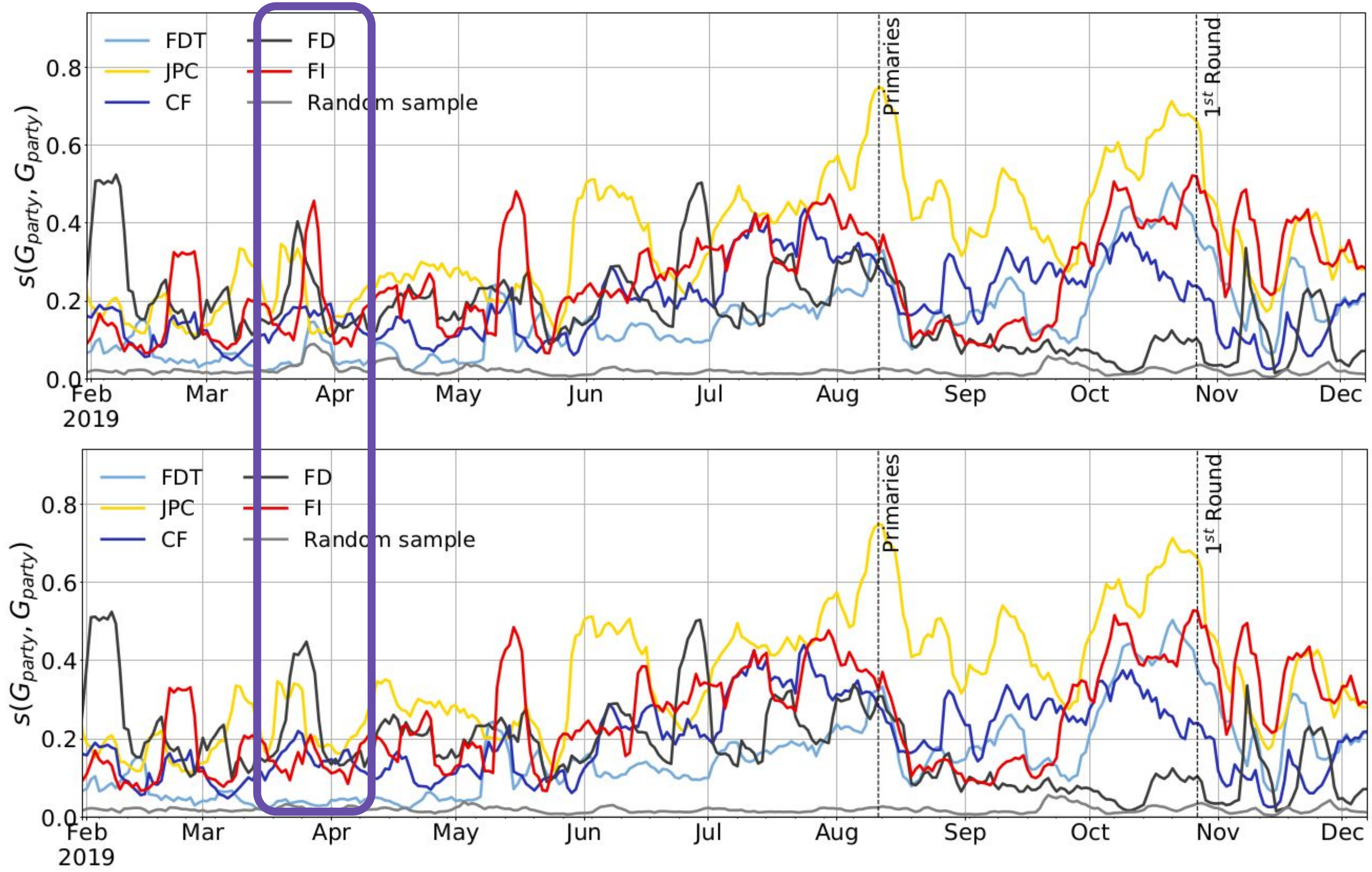


Center-left party (lost 1st round)

Center-left party (peronism)

Right party

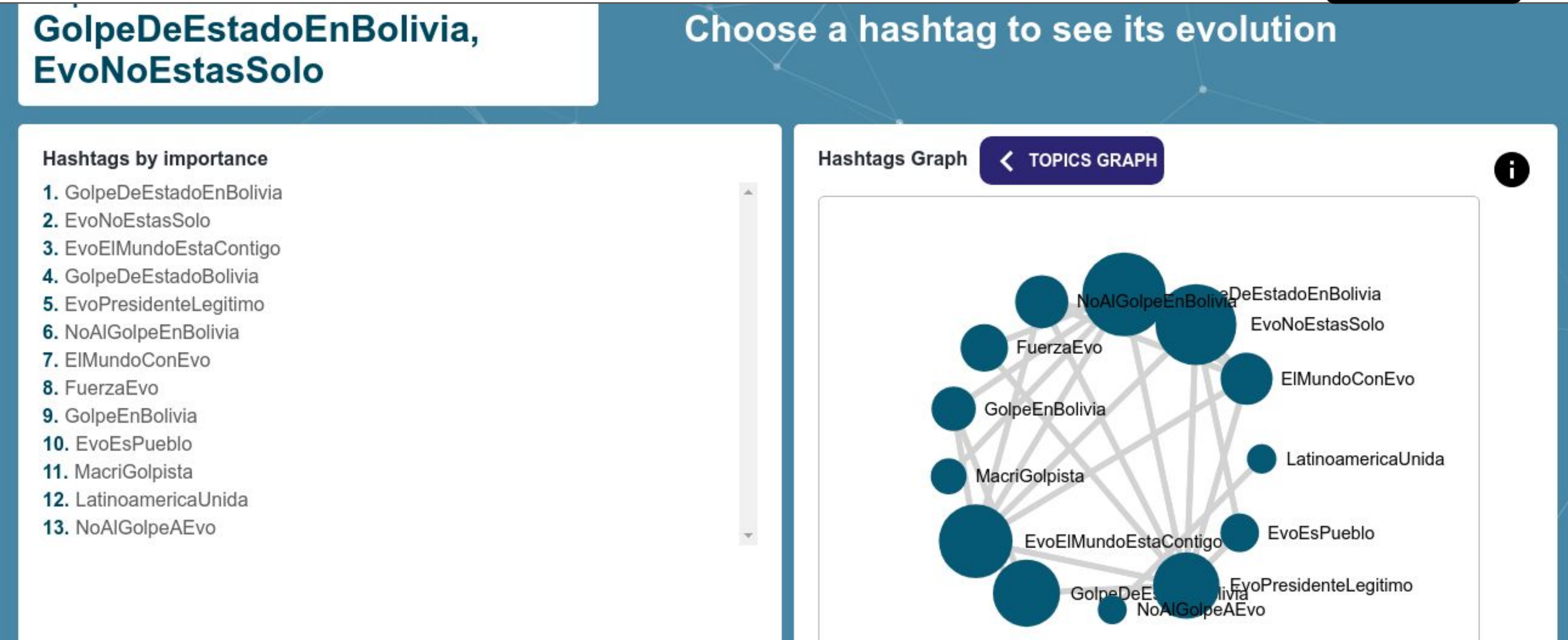
Result #3: By removing all hashtags from a specific topic and recomputing the similarities, we can correlate peaks in similarities with topics.



Robustness: We could reproduce our observations using 2 different community detection algorithms: OSLOM and Infomap.

Application: We deployed an online platform to track the topics' usage and the similarities between parties in real time. http://elecciones2019.fi.uba.ar/

SCAN ME



Thanks!

Acknowledgements:

CY ADVANCED STUDIES

Ministerio de Ciencia, Tecnología e Innovación
Argentina

anr

Trans-Atlantic Platform
Social Sciences and Humanities

DIGGING INTO DATA
CHALLENGE

Publication: Reyero, T. M., Beiró, M. G., Alvarez-Hamelin, J. I., Hernández, L., & Kotzinos, D. (2021). *Evolution of the political opinion landscape during electoral periods*. EPJ Data Science, 10(1), 31.

Result #2: Similarities can also reveal the reorientation of some supporters after a political event