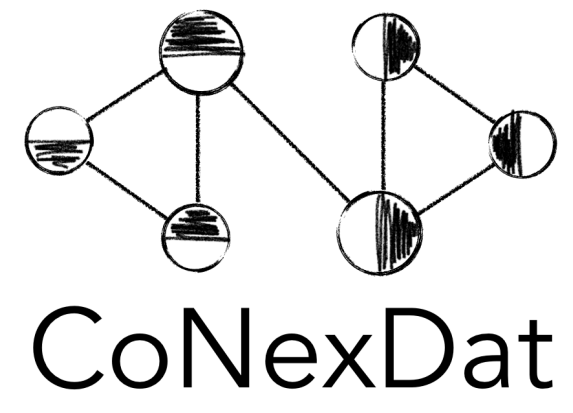


Topic-based study of a Twitter political network



CoNexDat

1



2



3



Tomás Mussi Reyero^(1,2)

Mariano G. Beiró^(1,2,3)

J. Ignacio Alvarez-Hamelin^(1,2,3)

OpLaDyn project



Introduction

- ◆ How to **extract** topics from a set Tweets?
- ◆ How can the **similarity** among groups of users be measured?

Dataset

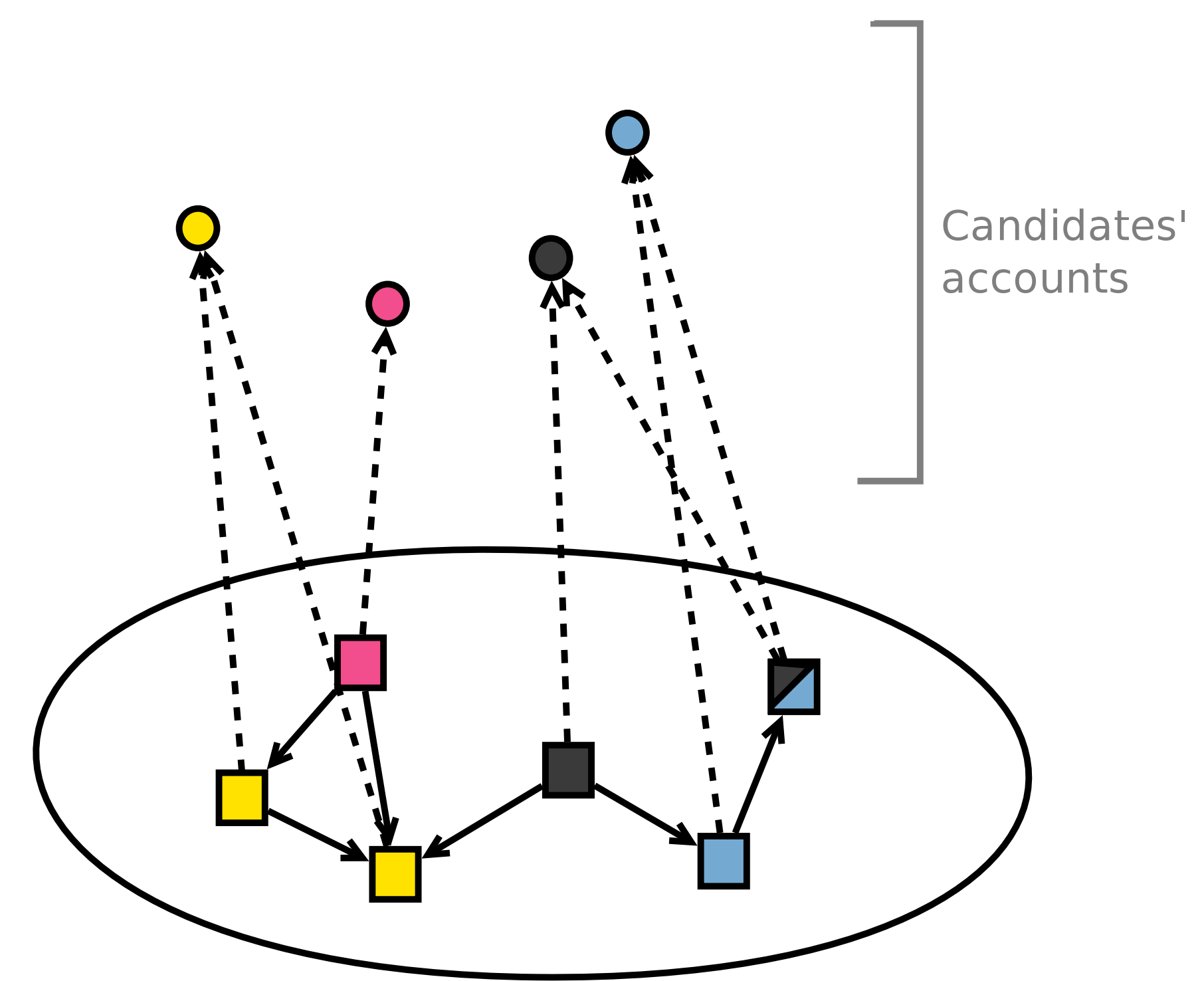
Twitter users following presidential candidates:

users: 300k (2015), 10M (2019)

candidates: 4 (2015), 5+5 (2019)

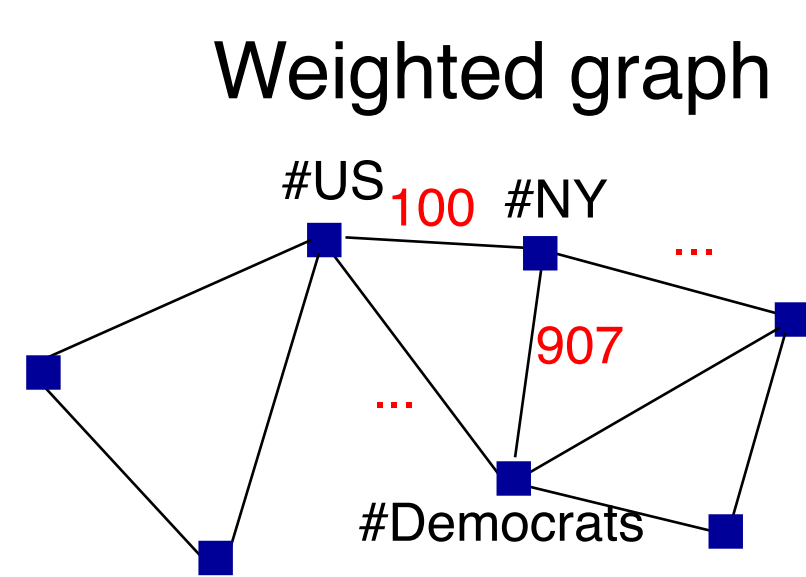
twits: 50M (2015), 900M (2019)

periods: 08/2015-03/2016, 01/2019, 11/2019

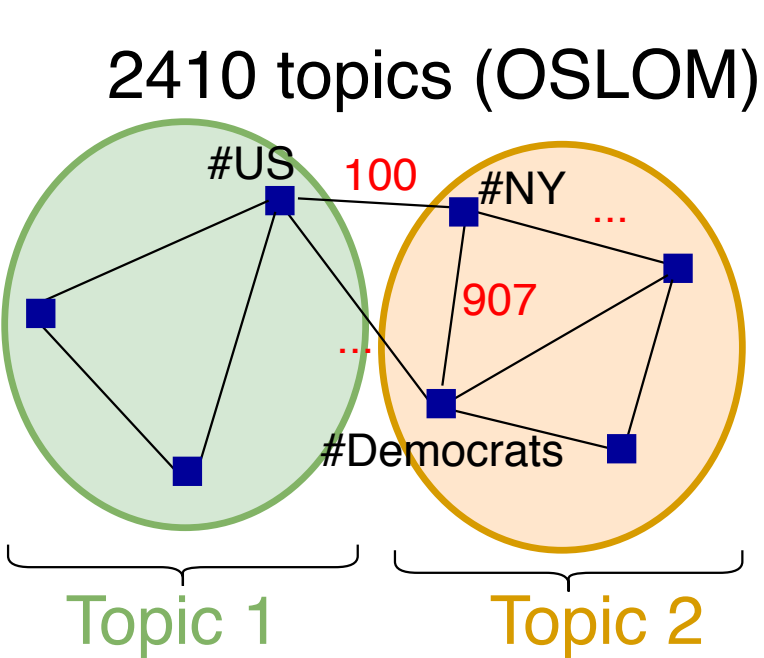


Methodology

hashtags co-occurrence



community detection

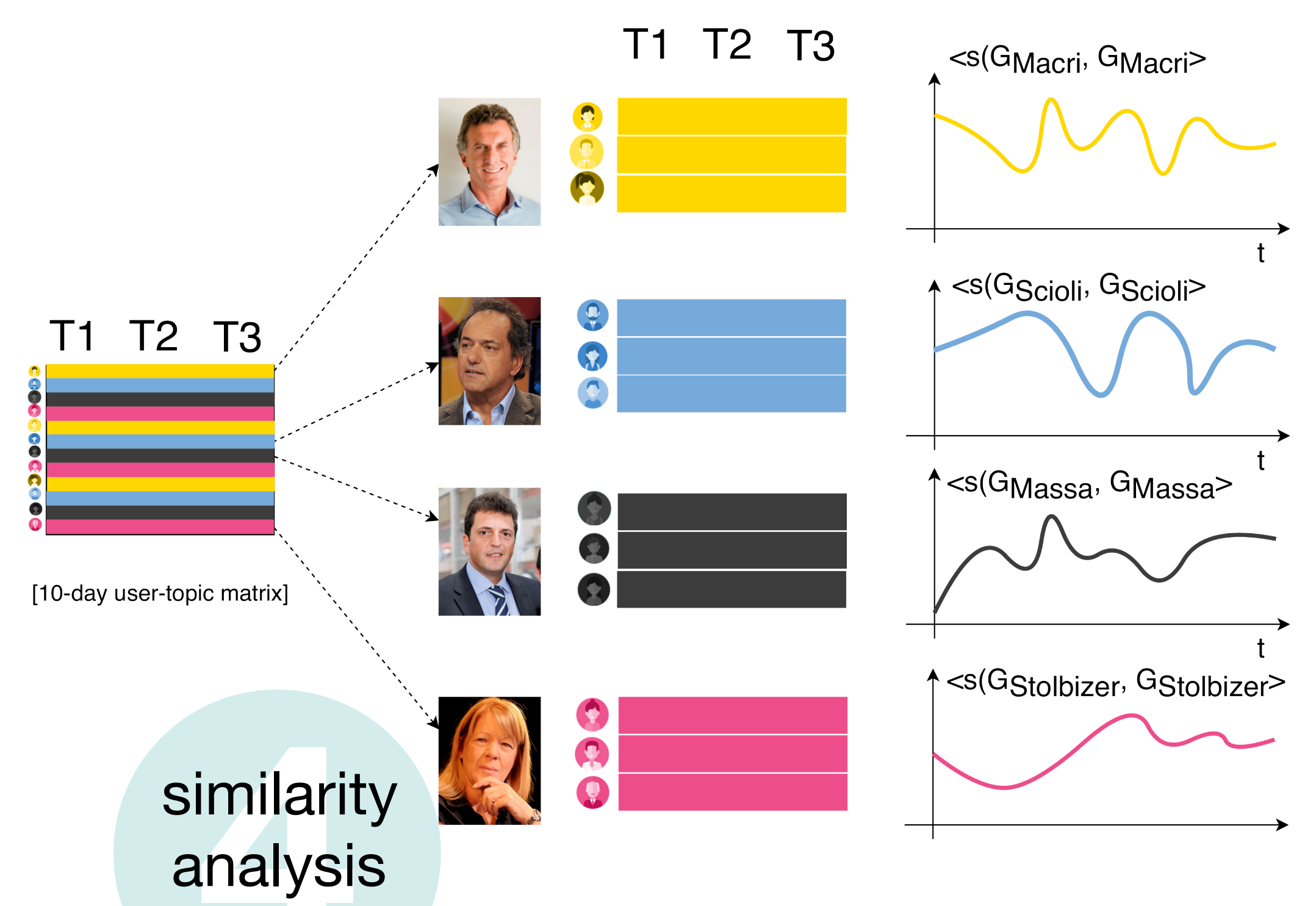


$$\text{TD-IDF} \& \frac{\vec{v}}{\|\vec{v}\|^2}$$

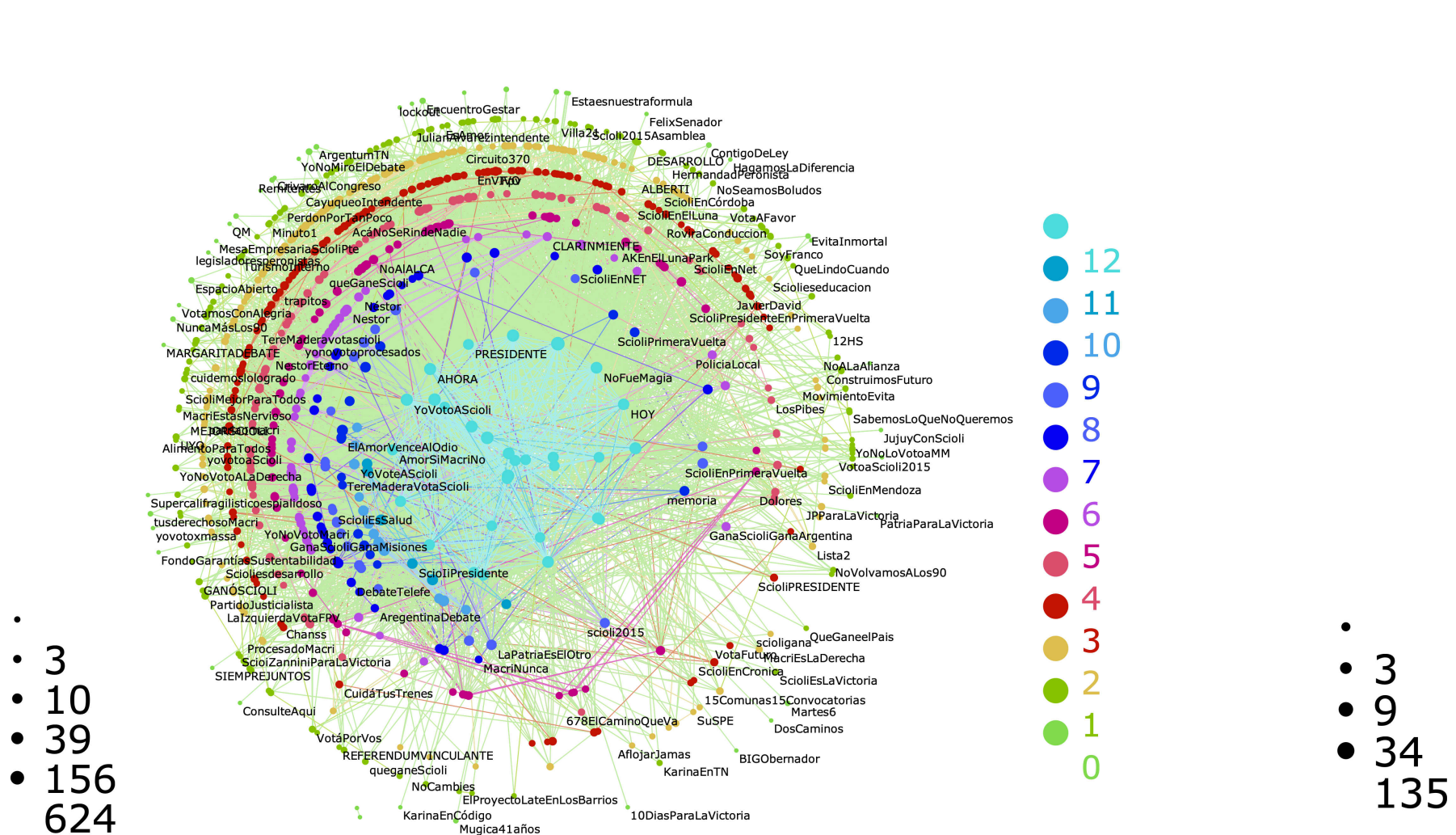
	Topic 1	Topic 2	Topic 3
#soyBean	0.303	0.043	0.952
#tractor	0.040	0.998	0.040
#NY	0.982	0.171	0.085
#Democrats			
#US first			
#GOP			

[user-topic matrix]

	T1	T2	T3
[user vector]	0.040	0.998	0.040



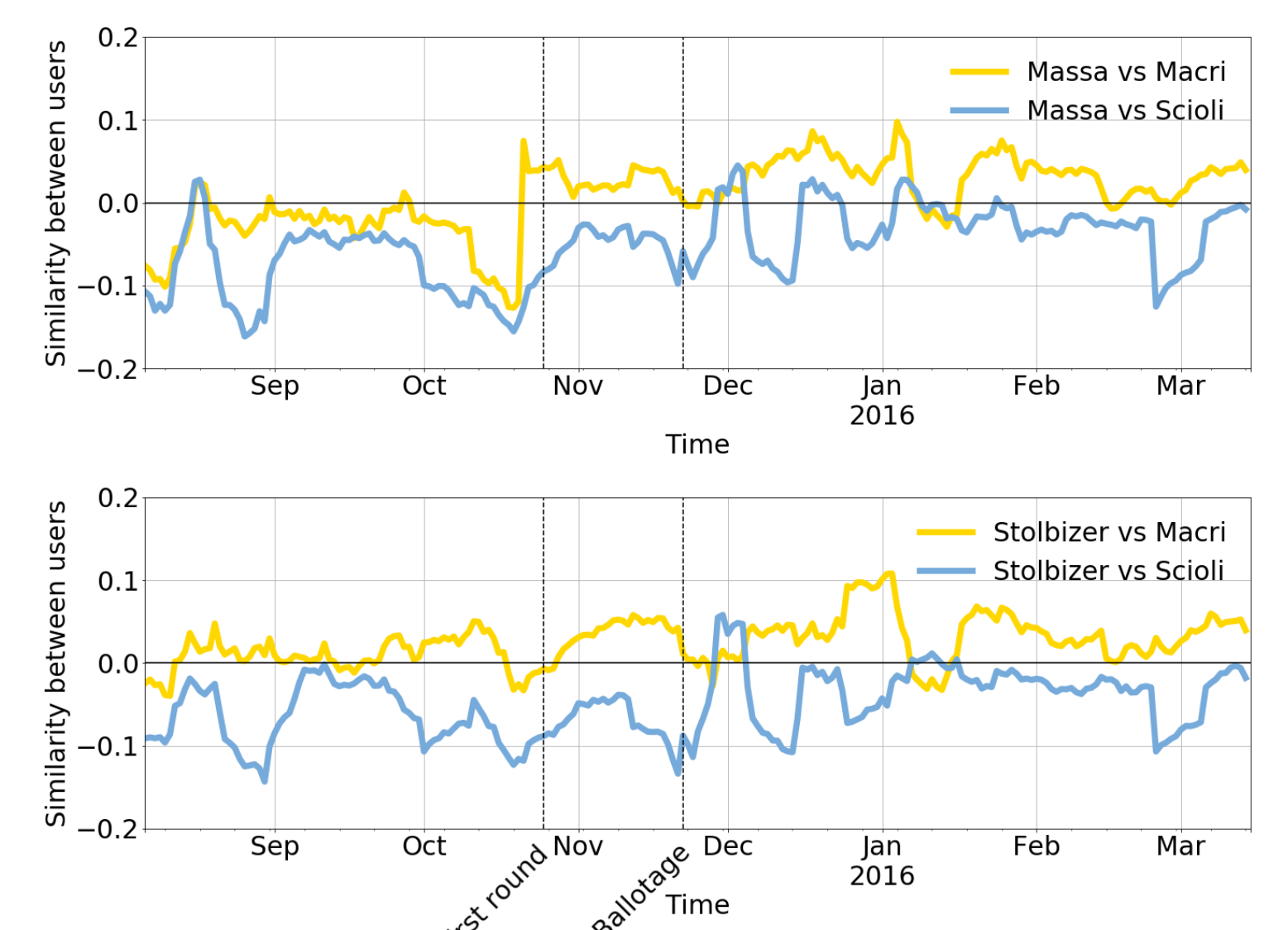
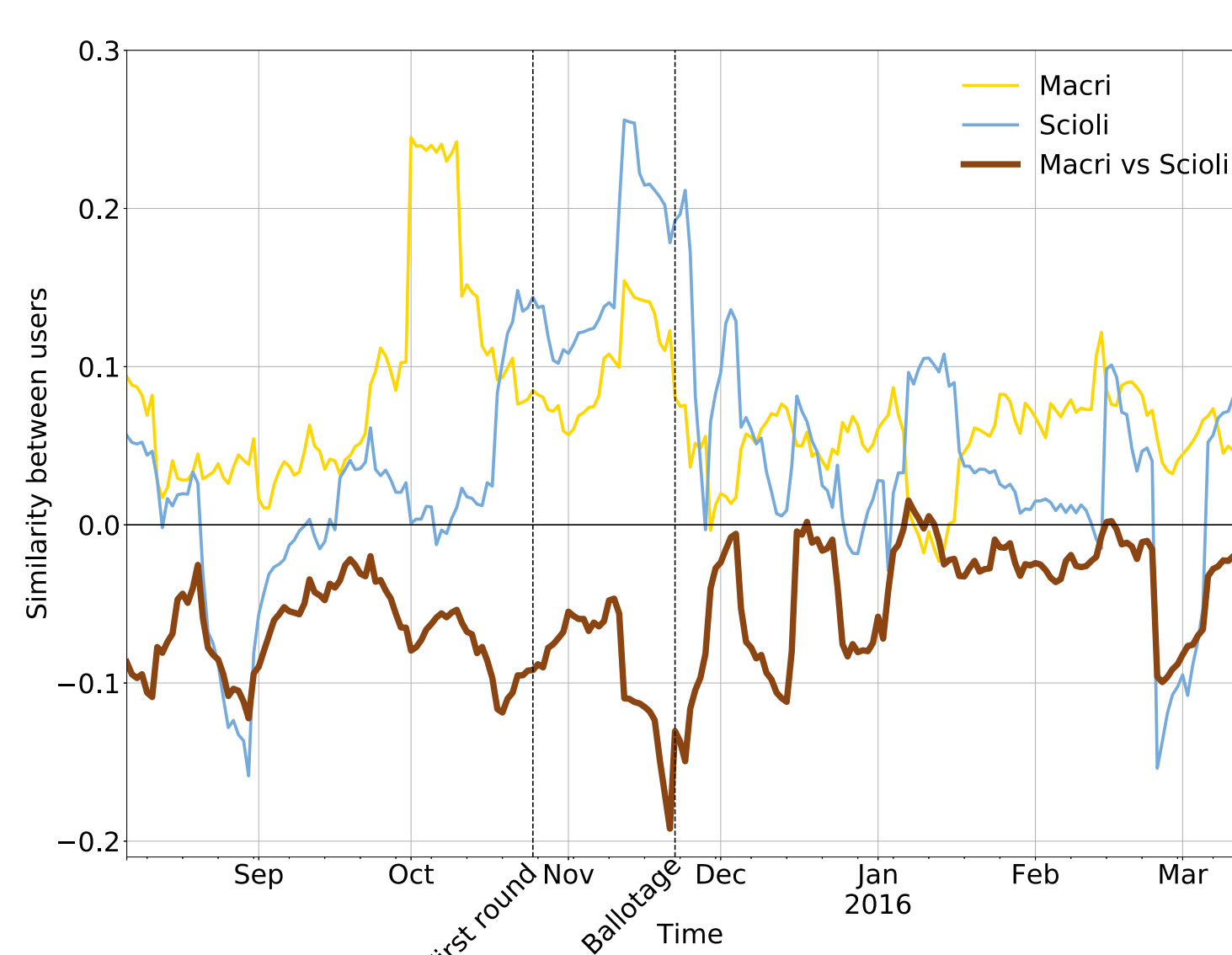
Results



K-dense decomposition for two topics

	2015	1st round	ballotage
Macri		34	51
Massa		21	-
Scioli		37	49
Stolbizer		3	-

Conclusion: Massa followers turn towards Macri



Test the result of the last Presidential election in Argentina

