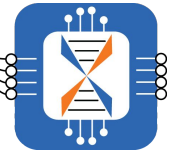
Communities of Protest

Jonathan Ronen,
Jennifer M. Larson,
Jonathan Nagler,
Joshua A. Tucker

Berlin, June 2017

NEW YORK UNIVERSITY

BIOINFORMATICS PLATFORM
MDC MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE BERLIN-BUCH
BIMSB THE BERLIN INSTITUTE FOR MEDICAL SYSTEMS BIOLOGY



About me

- 2005 - 2010 - MS.EE from Trondheim, Norway
- 2010 - 2013 - Software Developer in Oslo, Brussels, Munich
- 2013 - 2015 - Data Scientist, New York University
  - *This is where the work presented here was done*
- 2015 - Ph.D. Fellow, Max Delbruck Center, Berlin

@jonathanronen

yablee@gmail.com

Place de la Republique, January 11th 2015

Protests are social

- Conventional wisdom: the decision to protest is social.

- But little empirical evidence of the importance of social ties and network structure.

Post protest surveys

Ask participants if they knew other protesters

Opp and Gern, 1993

McAdam and Paulsen, 1993

Post protest surveys

Ask participants if they knew other protesters

- Retrospective
- Small N
- Incomplete ties to protesters
- Often no ties to non-protesters
- No network structure

Opp and Gern, 1993

McAdam and Paulsen, 1993

Social media for empirical data

Twitter networks in social science

- Birds of the Same Feather Tweet Together. Bayesian Ideal Point Estimation Using Twitter Data. (Barbera, 2015)

- Sectarian Twitter Wars: Sunni-Shia Conflict and Cooperation in the Digital Age. (Siegel, 2015)

Wish list

- A set of confirmed protesters
- Their social ties, including to other protesters and to non-protesters
- The ties of their ties - network structure
- A control set of non protesters and their network structure

Collection of the dataset

Before the march (after the attack), we set up a stream collecting some of the twitter conversation around Charlie Hebdo

```
POST https://stream.twitter.com/1.1/statuses/filter.json

    {

        “track”: ["CharlieHebdo", "JeSuisCharlie", "Charlie Hebdo",

                  "JeSuisAhmed", "JeNeSuisPasCharlie", "#Beinfait",

                  "#JeSuisKouachi"]

    }
```

Collection of the dataset *(cont.)*

- Before the march (after the attack), we set up a stream collecting some of the twitter conversation around Charlie Hebdo

- Identify 764 users who tweeted with geotagging from the Place de la Republique during protest hours

Collection of the dataset *(cont.)*

- Before the march (after the attack), we set up a stream collecting some of the twitter conversation around Charlie Hebdo

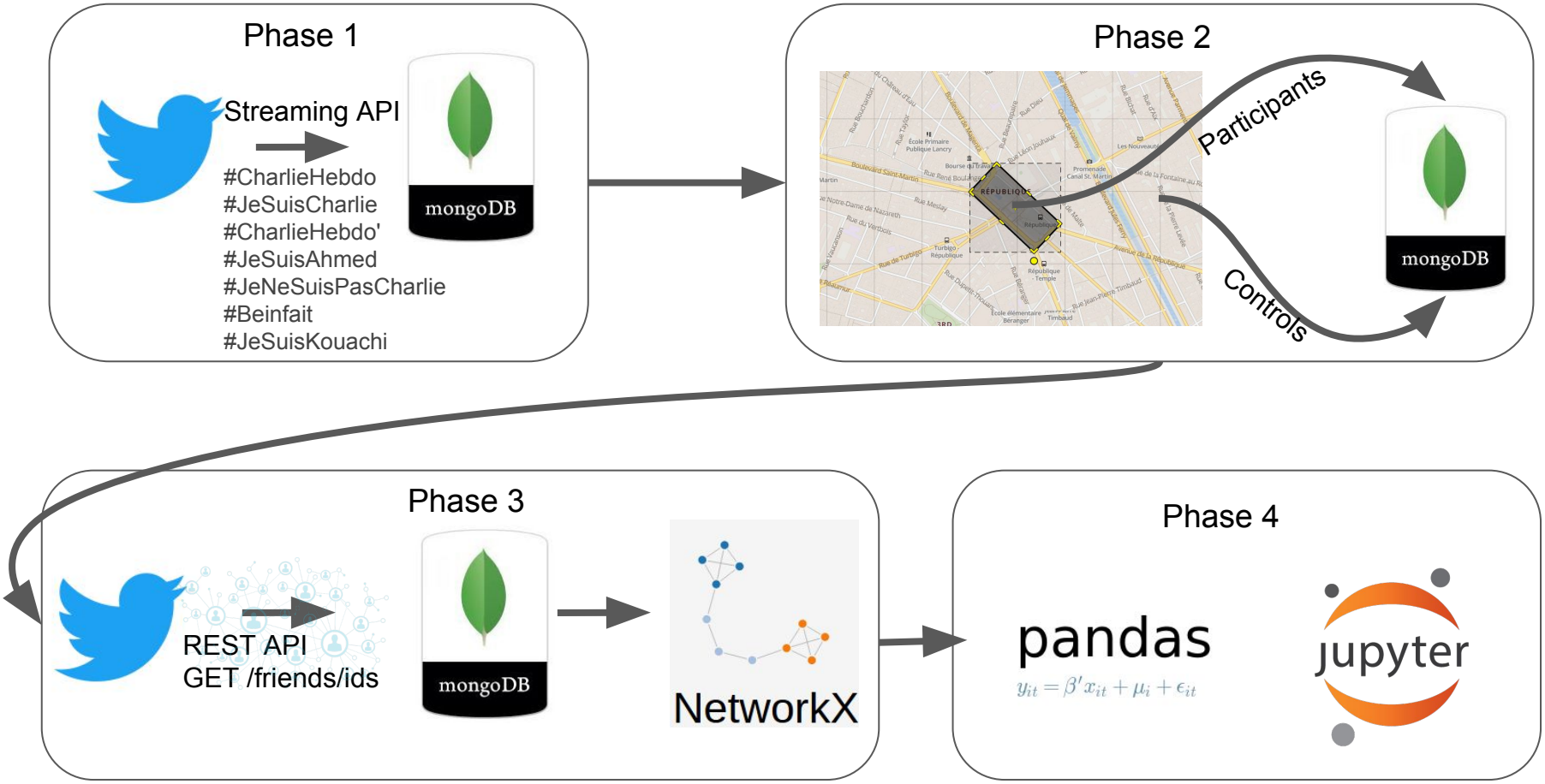
- Identify 764 users who tweeted with geotagging from the Place de la Republique during protest hours

- Identify 1,000 control users who tweeted *a distance away from* the march, at the time
  - This is no guarantee they did not attend, or intend to attend, and that only makes any further claims more robust

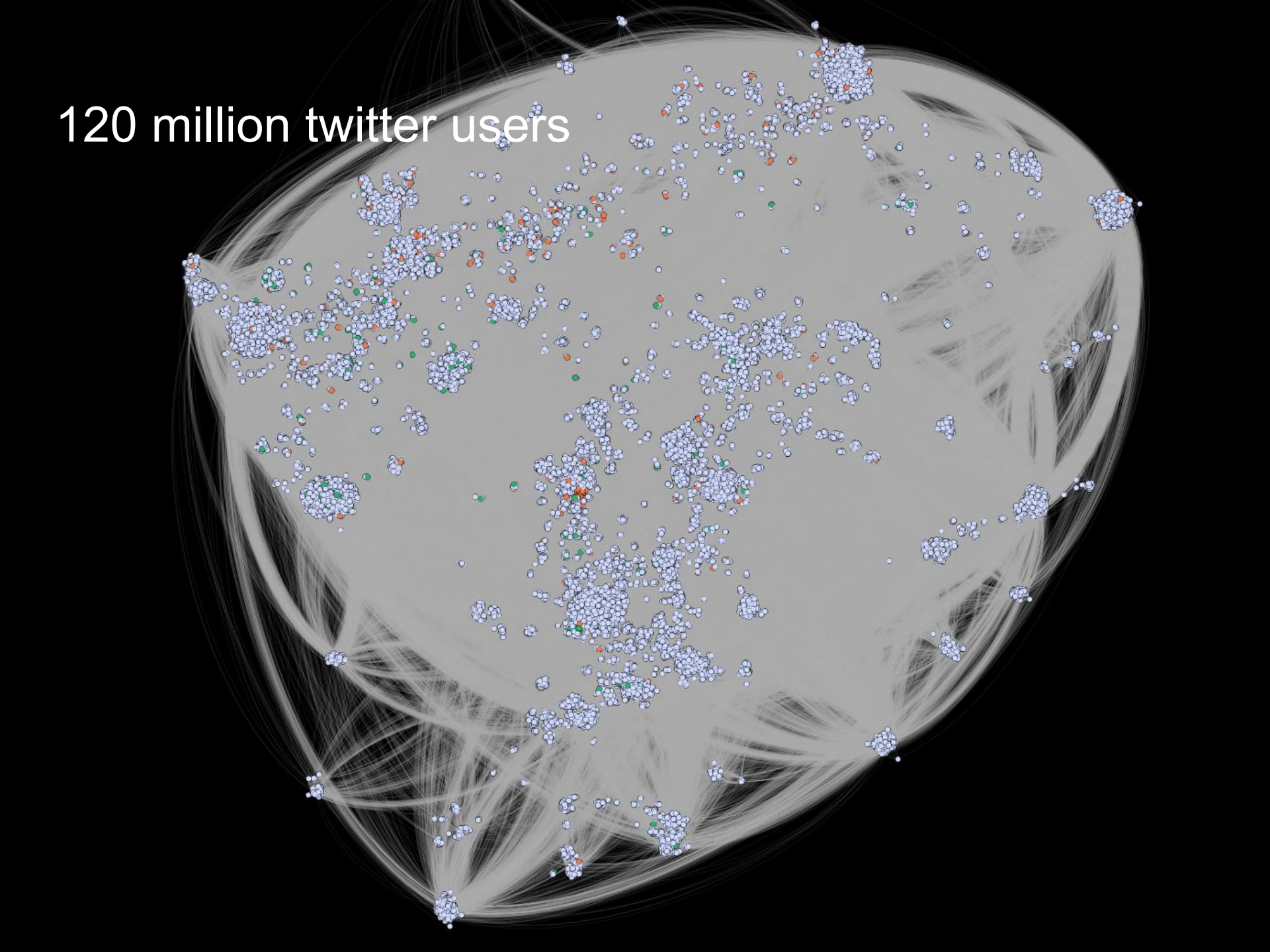
Collection of the dataset *(cont.)*

- Before the march (after the attack), we set up a stream collecting some of the twitter conversation around Charlie Hebdo
- Identify 764 users who tweeted with geotagging from the Place de la Republique during protest hours
- Identify 1,000 control users who tweeted *a distance away from* the march, at the time
  - This is no guarantee they did not attend, or intend to attend, and that only makes any further claims more robust
- Build out a 2-degree network
  - Friends, and friends of friends

Analysis pipeline



120 million twitter users



Our Approach: A Network Theory of Protest

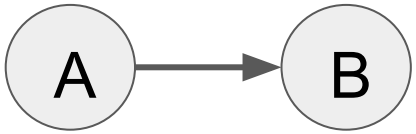
- Individuals protest if they value the protest highly enough.

- Exposure to other individuals who value the protest highly, increases individuals' own valuation of protest.

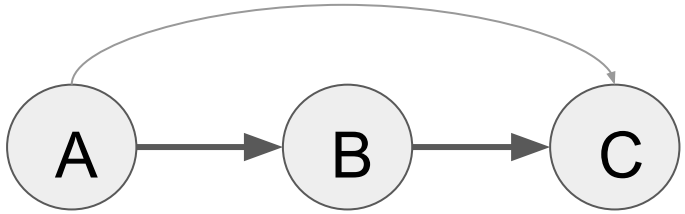
- Exposure on twitter is a function of network distance and tie strength.

A Network Theory of Protest: definitions

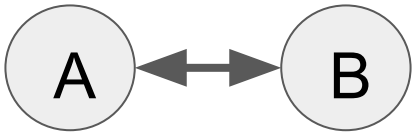
- Tie: A follows B
  - A has a tie to B



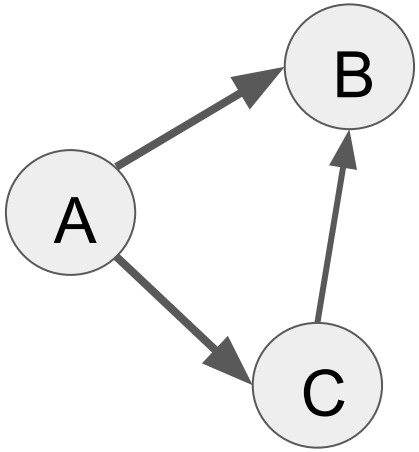
- Ties tie: A follows B who follows C
  - A has a ties-tie to C



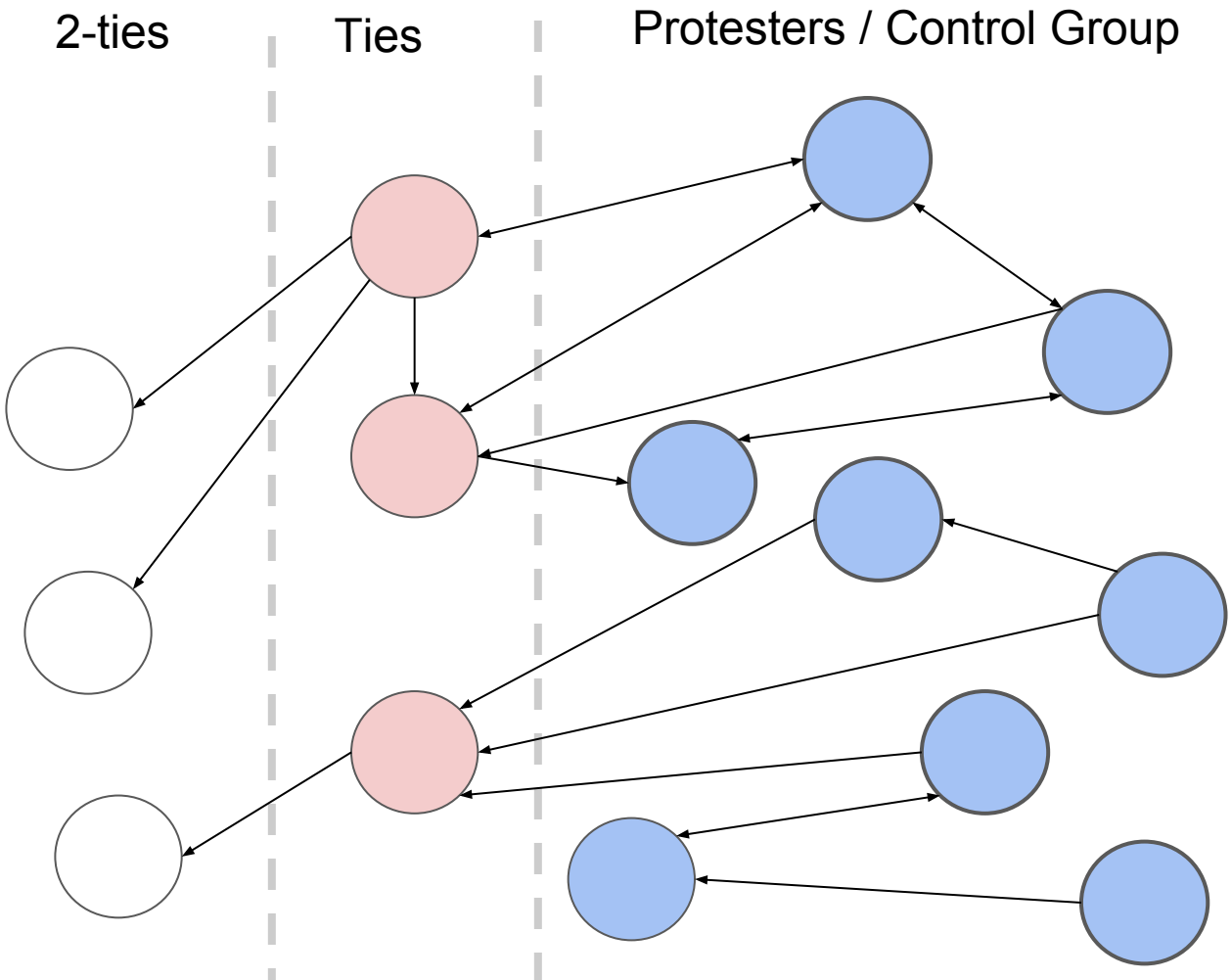
- Strong (reciprocated) tie: follow each other
  - A and B have a strong tie



- Triad
  - Follows two people, at least one follows the other

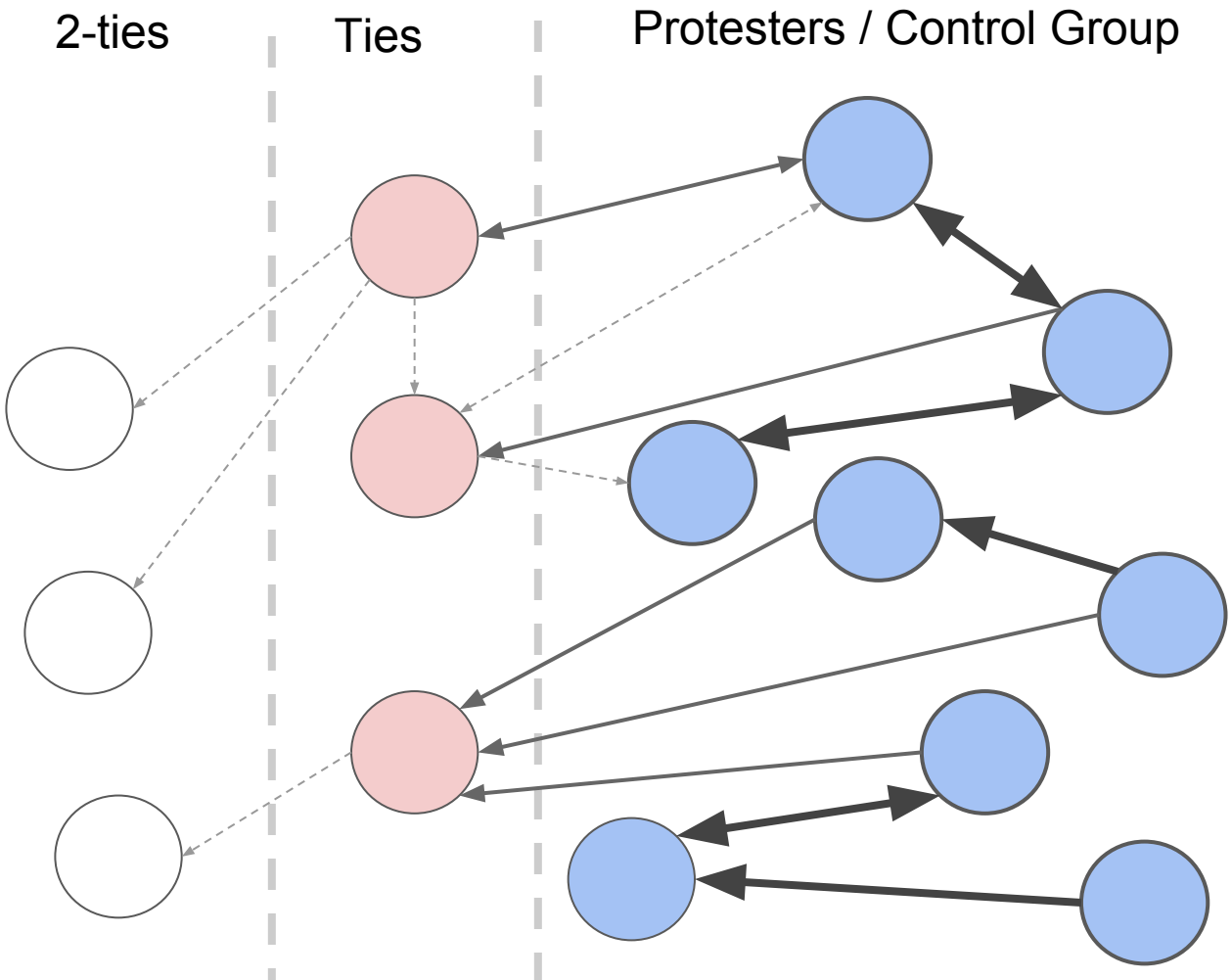


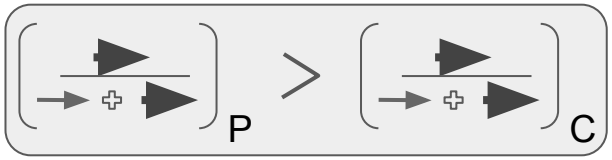
A Twitter Follower Network



Hypothesis 1: Protesters have more ties to other protesters

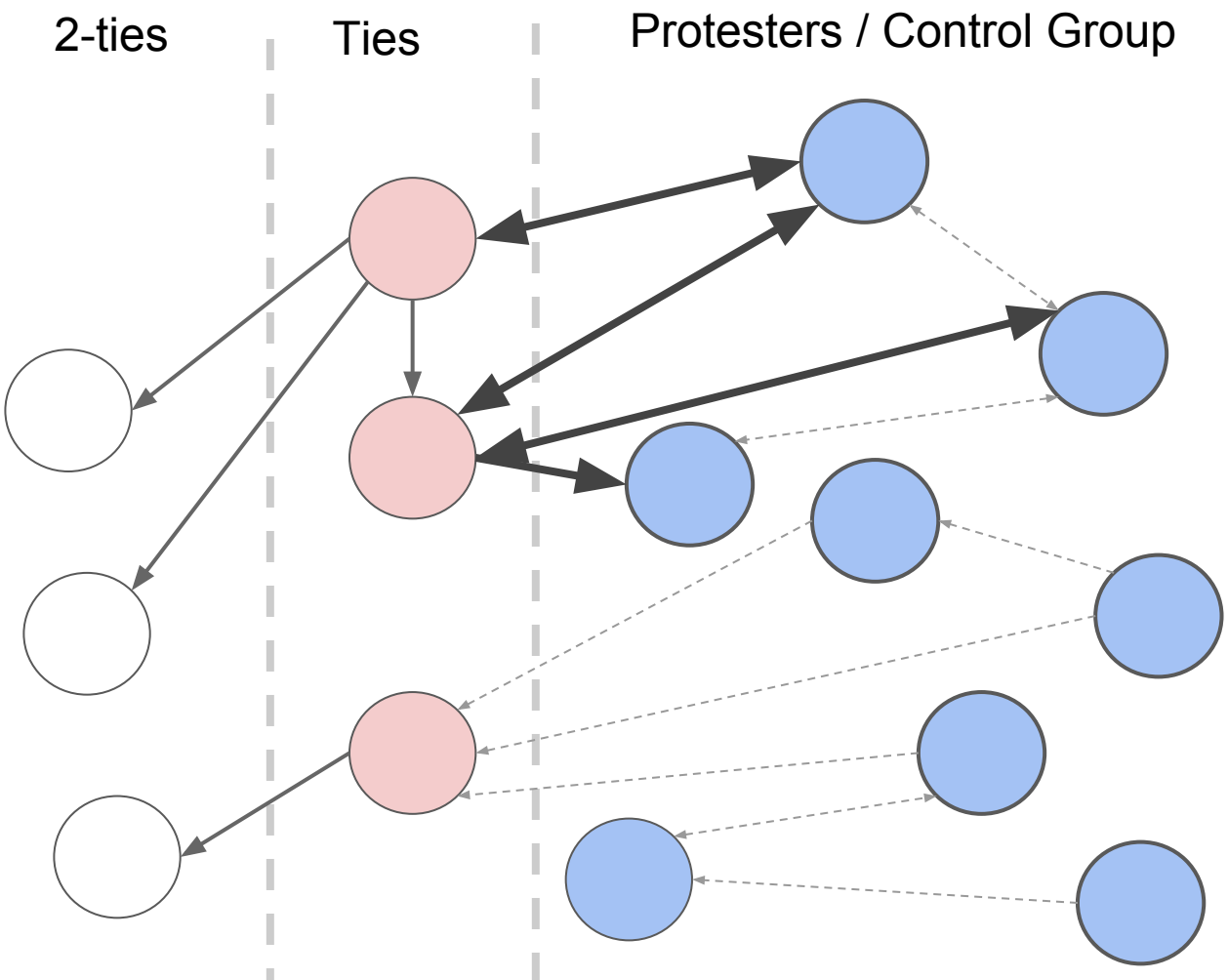
We expect the proportion of protesters' ties that go to other protesters to be higher than for the control group.

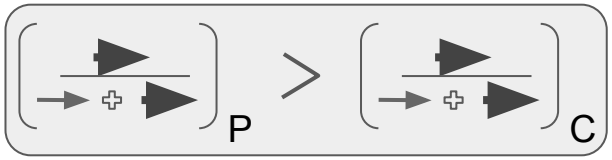




Hypothesis 2: Protesters have more ties-of-ties to each other

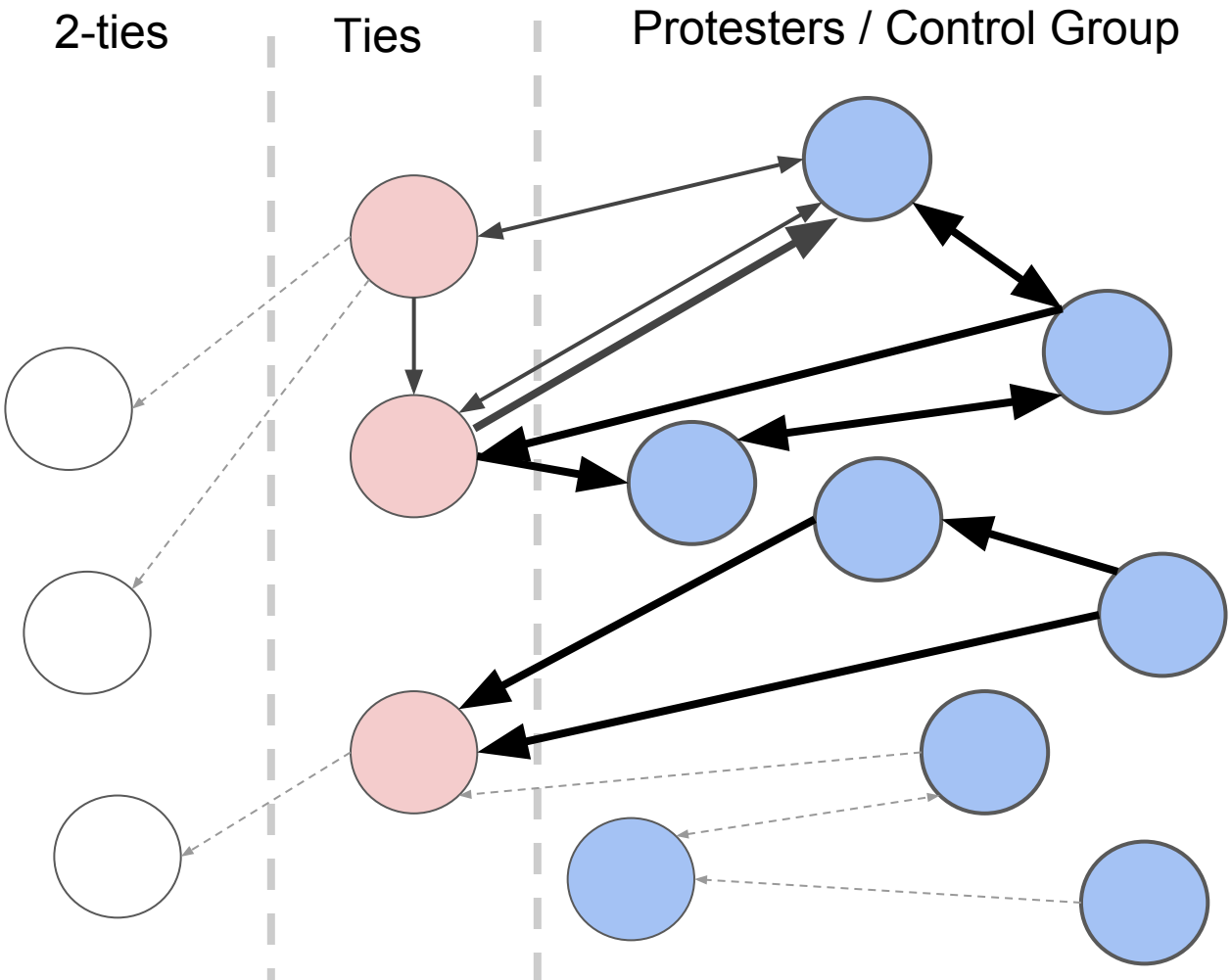
We expect the proportion of protesters' ties-of-ties that go back to other protesters to be higher than the control.

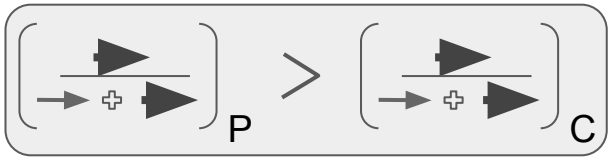




Hypothesis 3: Protesters have more triads with at least one protester, than controls do with controls

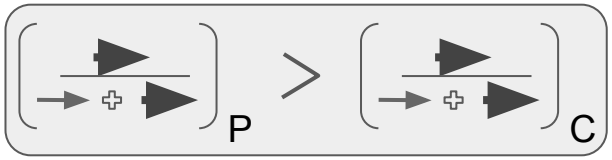
We expect protesters to form more triads with other protesters, than controls with other controls.

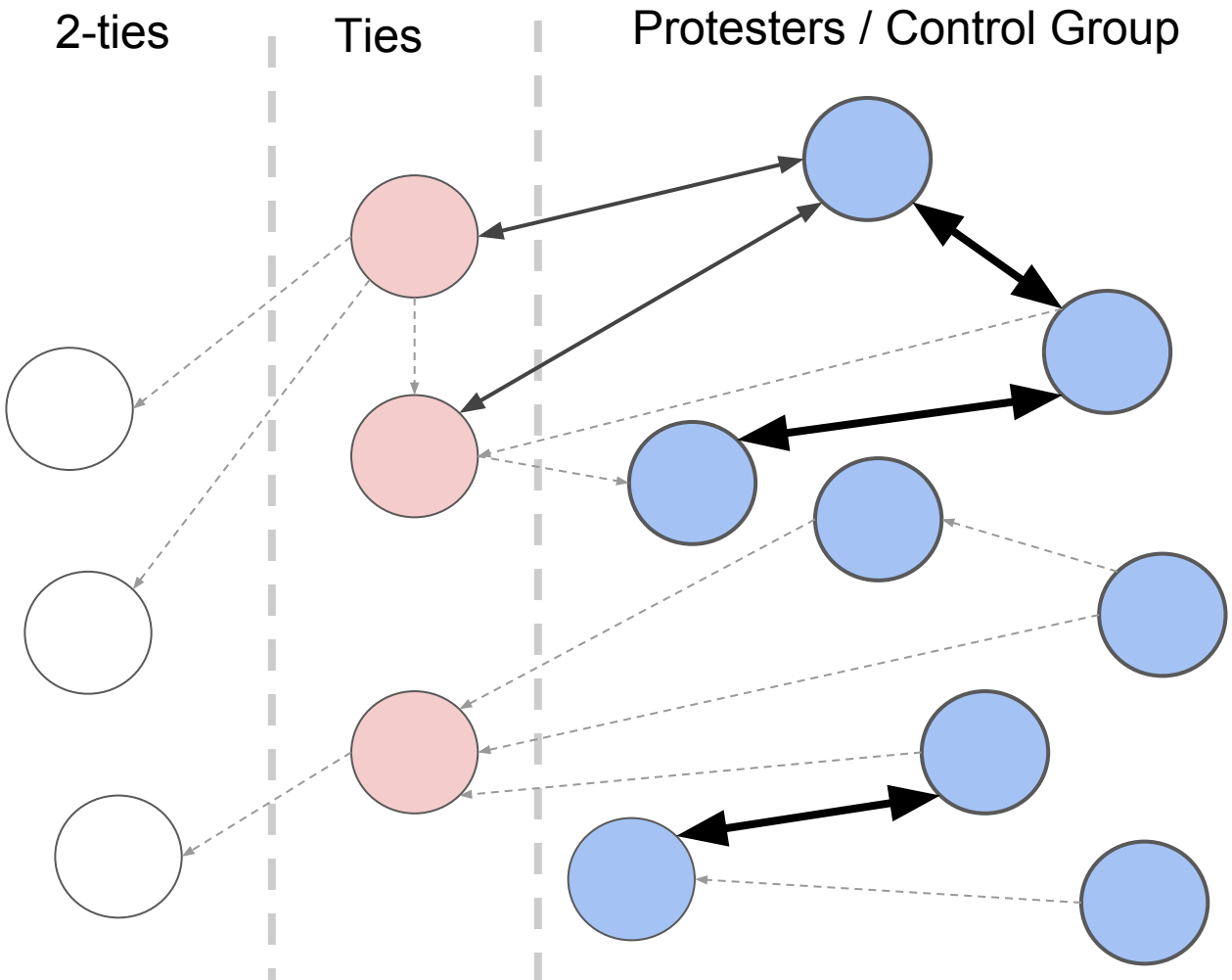




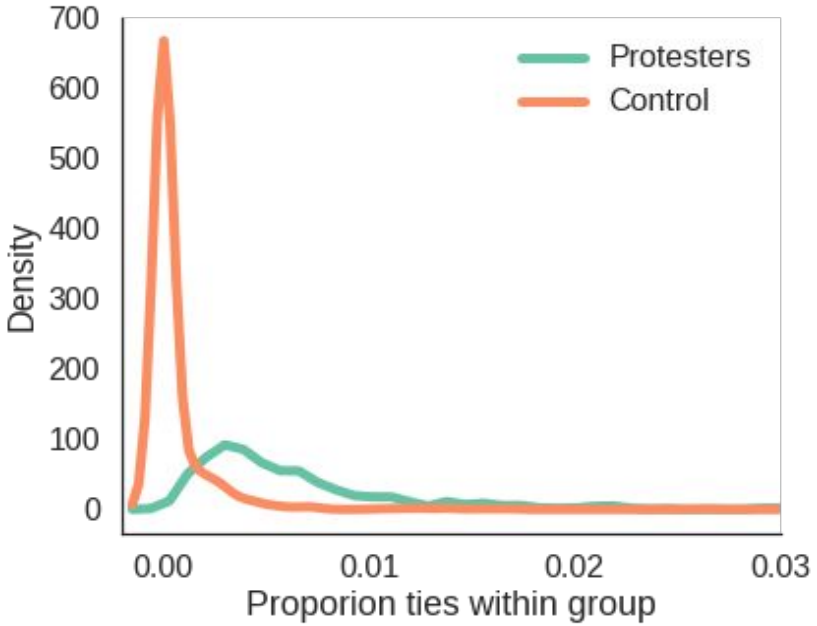
Hypothesis 4: Protesters have more strong ties to other protesters

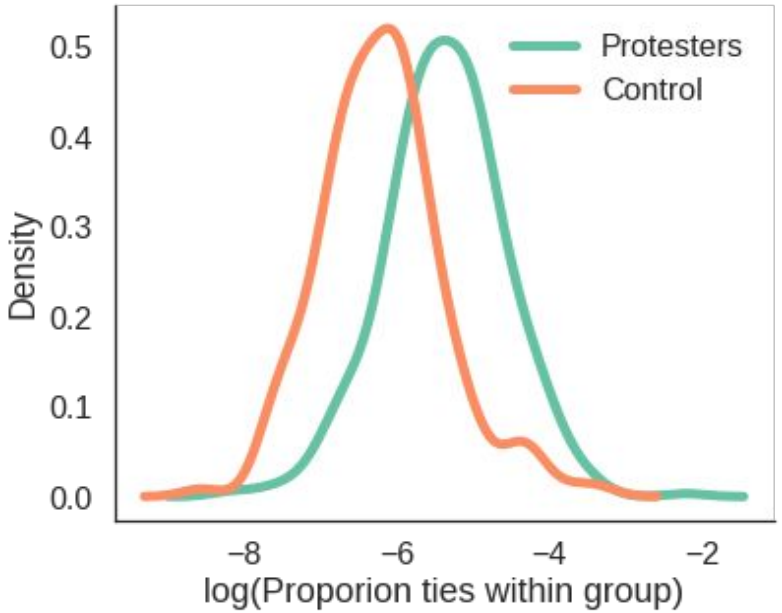
We expect protesters' proportion of strong (reciprocated) ties to each other to be higher than controls to controls.





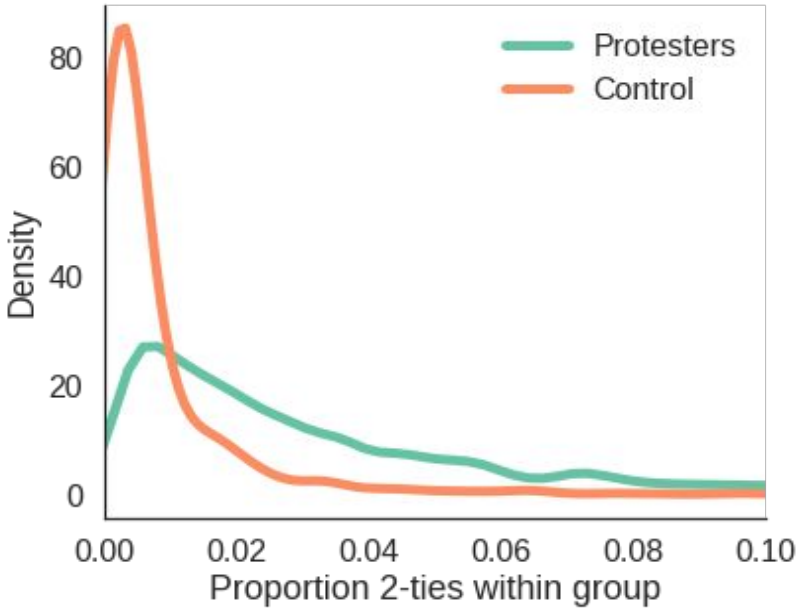
Results - Hypothesis 1 (ties within group)

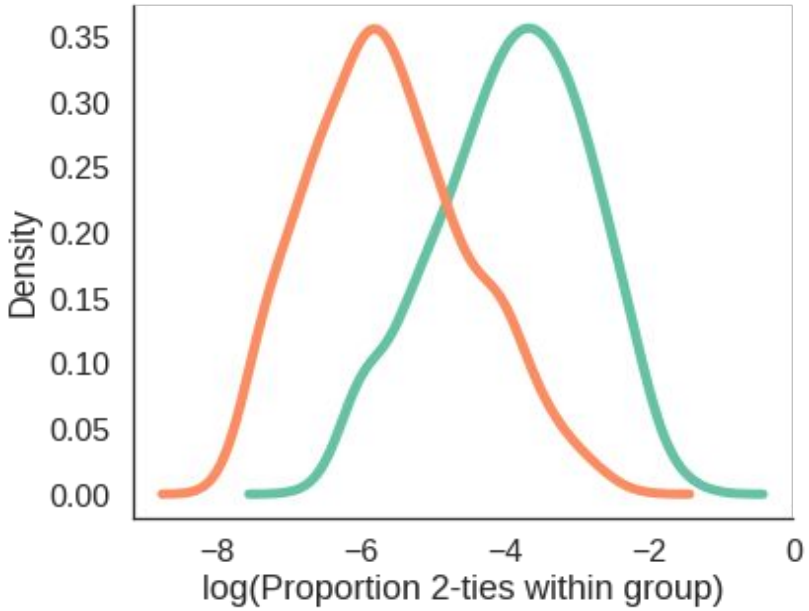




Protesters	Control	T statistic	P value
0.00400 +/- 0.00633	0.00057 +/- 0.00210	15.91171 [log: 23.82256]	2.39177E-53 [log: 6.93E-109]

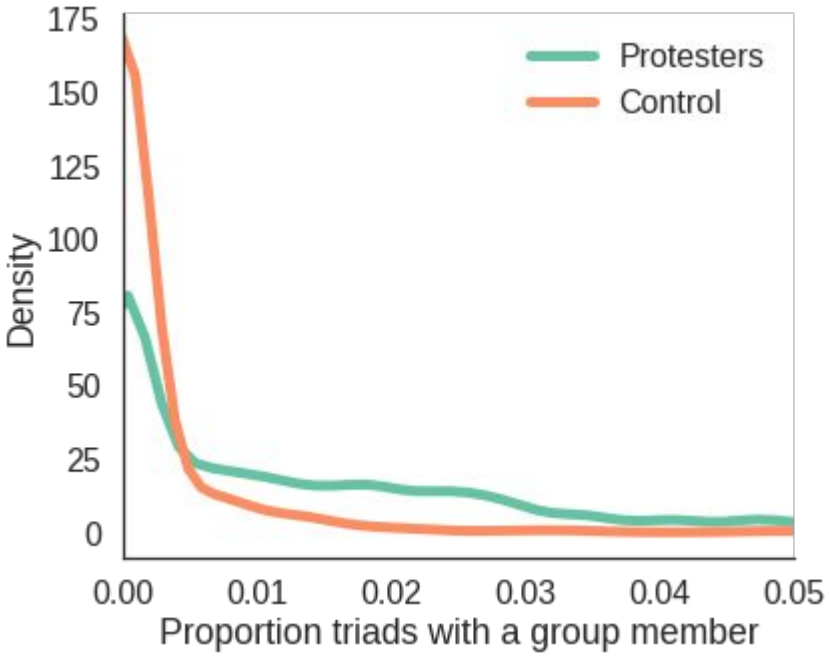
Results - Hypothesis 2 (ties of ties within group)

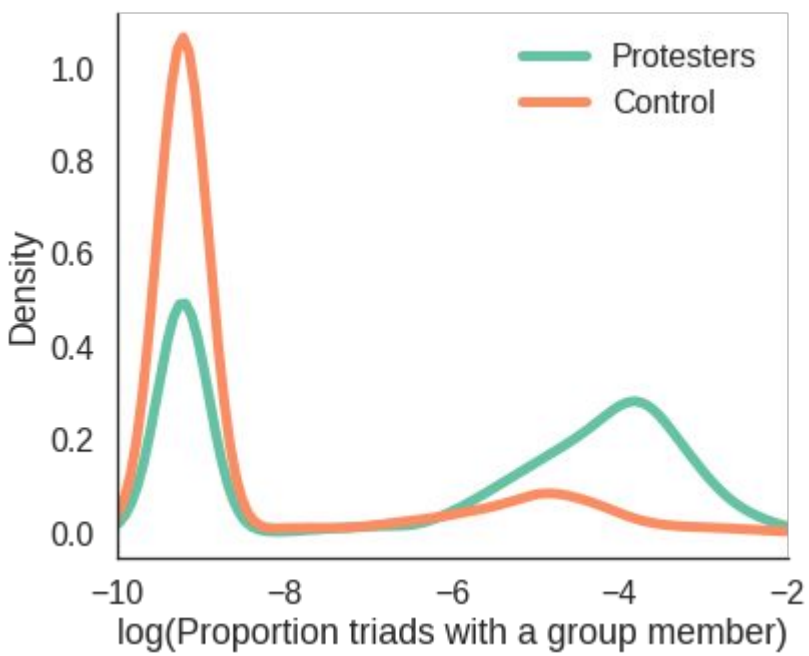




Protesters	Control	T statistic	P value
0.03260 +/- 0.03275	0.00706 +/- 0.01020	23.12424 [log: 32.91035]	1.84286E-103 [log: 2.50E-185]

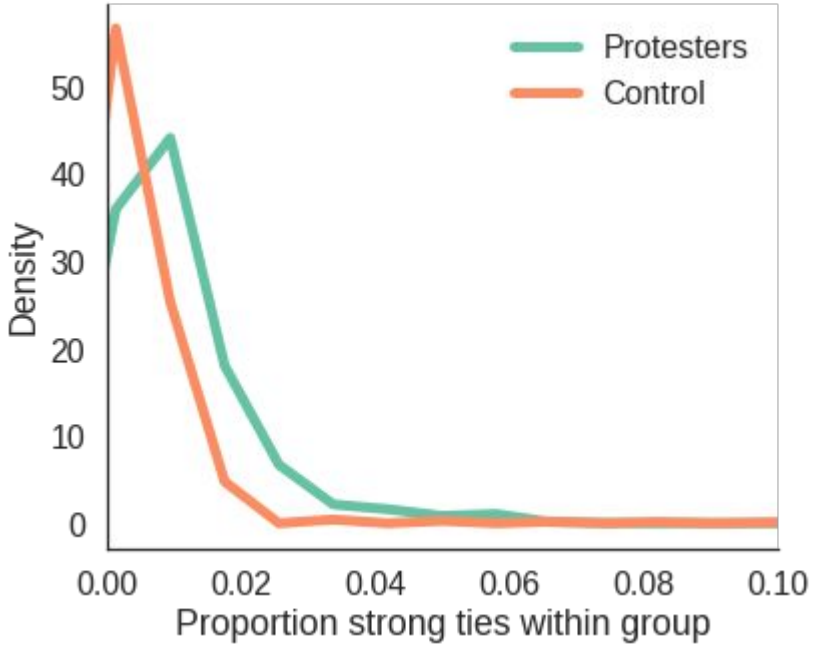
Results - Hypothesis 3 (triads with another protester)

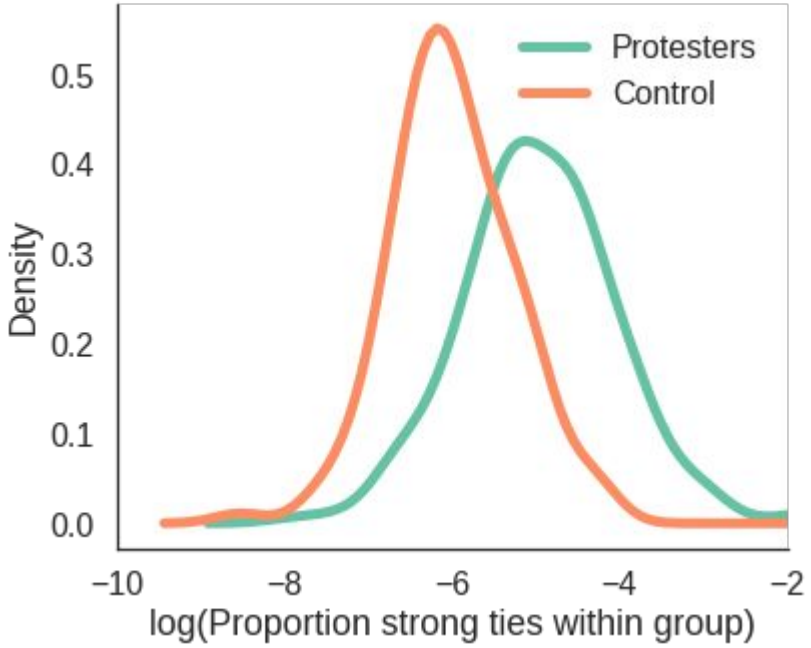




Protesters	Control	T statistic	P value
0.01377 +/- 0.01898	0.00213 +/- 0.00798	17.08805 [log: 22.57068]	1.46802E-60 [log: 6.19E-99]

Results - Hypothesis 4 (strong ties)





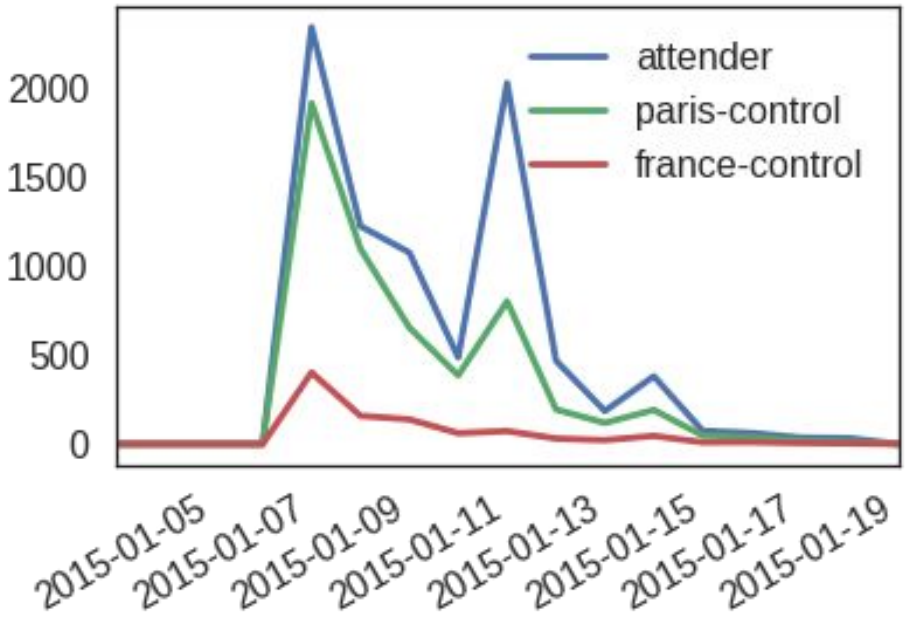
Protesters	Control	T statistic	P value
0.01973 +/- 0.09206	0.01010 +/- 0.08492	1.06702 [log: 10.54028]	2.86483E-01 [log: 1.40E-23]

Robustness

- All tests robust to
  - France control set
    - Similar to Paris in all accounts
  - Removing verified accounts
- Place de la Republique is not often geotweeted from
  - Collected tweets from a geobox for 2 weeks - 0 results

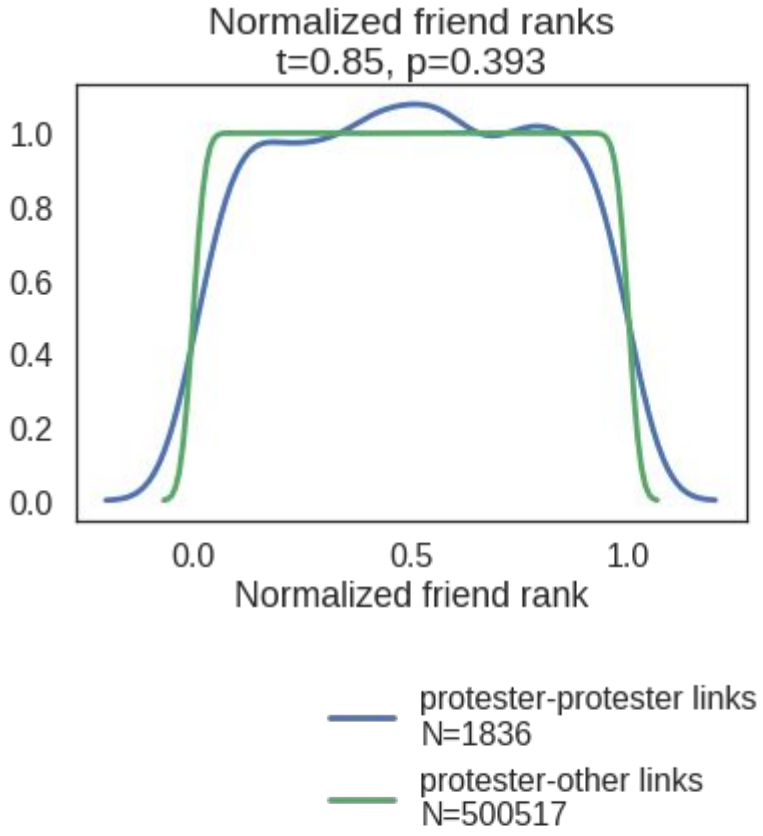
Robustness

- All tests robust to
  - France control set
    - Similar to Paris in all accounts
  - Removing verified accounts
- Place de la Republique is not often geotweeted from
  - Collected tweets from a geobox for 2 weeks - 0 results
- Interest in the charlie hebdo massacre.



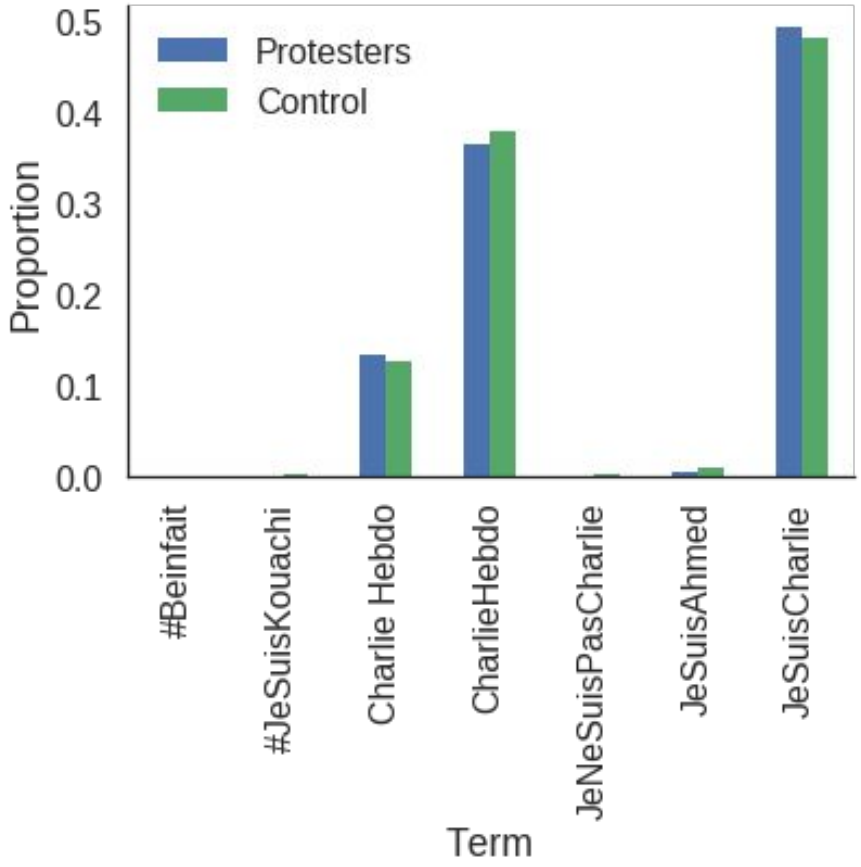
Robustness

- All tests robust to
  - France control set
    - Similar to Paris in all accounts
  - Removing verified accounts
- Place de la Republique is not often geotweeted from
  - Collected tweets from a geobox for 2 weeks - 0 results
- Interest in the charlie hebdo massacre.
- Friendships formed before protest.



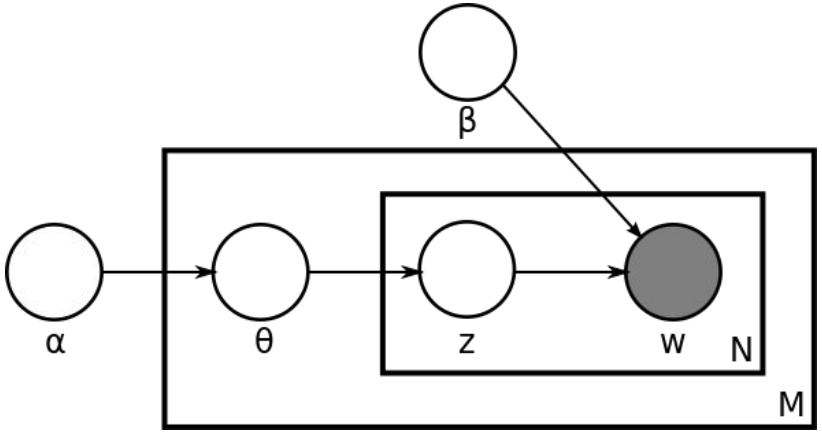
Robustness

- All tests robust to
  - France control set
    - Similar to Paris in all accounts
  - Removing verified accounts
- Place de la Republique is not often geotweeted from
  - Collected tweets from a geobox for 2 weeks - 0 results
- Interest in the charlie hebdo massacre.
- Friendships formed before protest
- #JeNeSuisPasCharlie

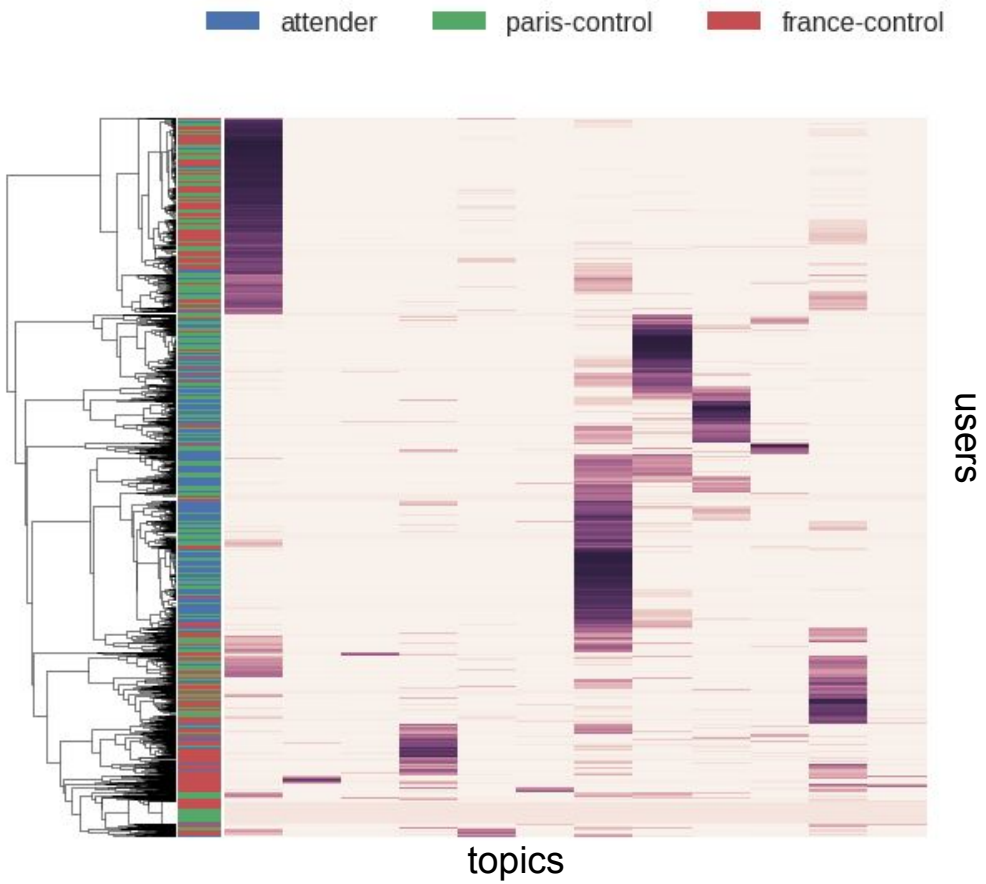


Topic modeling the friends lists

Latent Dirichlet Association



- Look at the most followed accounts
- When following someone, a user first picks a “topic”, and then picks an account from that “topic”
- Users have their own topic distributions
- Topics have their own account distributions



Conclusions

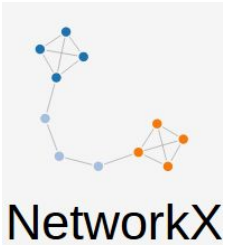
- Evidence for network structure hypotheses (H1, H2)
- Evidence for tie strength hypotheses (H3, H4)
- Robust to alternative control groups, removing verified accounts.
- Evidence consistent with protest decision as network function
- Door is now open to social networks research

Outlook

- Observational data - what can we do to add causality?

- Can we use the tweet texts?

A big thank you to the tool makers



Acknowledgments

Jennifer M. Larson

Jonathan Nagler

Joshua Tucker

The Social Media and Political Participation lab, New York University

- http://smapp.nyu.edu