#DebateNight: The Role and Influence of **Socialbots** in the Democratic Process

Marian-Andrei Rizoiu Timothy Graham Rui Zhang

Yifei Zhang Robert Ackland Lexing Xie

ComputationalMedia @ANU: http://cm.cecs.anu.edu.au

Two influencers: the 2016 U.S. Presidential elections

60k followers

136k followers

Common traits:

- Pro-republican;
- Highly influential, highly followed and retweeted;
- Opinion leaders;
- ...

Two influencers: the 2016 U.S. Presidential elections

60k followers

136k followers

Common traits:

- Pro-republican;
- Highly influential, highly followed and retweeted;
- Opinion leaders;
- ...

Russian-controlled bots operated by the Internet Research Agency in St. Petersburg

[Forbes, The Guardian, CNN + 50 more]

The political influence of socialbots

SocialBots:

"Software processes that are programmed to appear to be human-generated within the context of social networking sites such as Facebook and Twitter"

(Gehl and Bakardjieva 2016, p.2)

Immediate and long term research questions:

- are socialbots influential in the political discourse?
- did they have political partisanship?
- *(long term)* were they instrumental for the results of the elections?

#DebateNight dataset

- First U.S. Presidential Debate
  (26 sept 2016, 8.45pm to 10.45pm EDT)
- Twitter Firehose

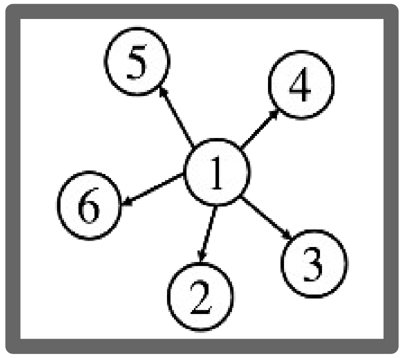
Dataset stats:

- length: **90 minutes**
- #tweets: **6.5M**
- #users: **1.45M**

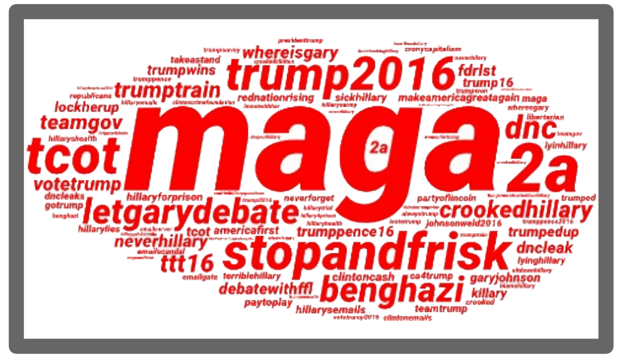
Hashtags:

#DebateNight
#Debates2016
#election2016
#HillaryClinton
#Debates,
#Hillary2016
#DonaldTrump
#Trump2016

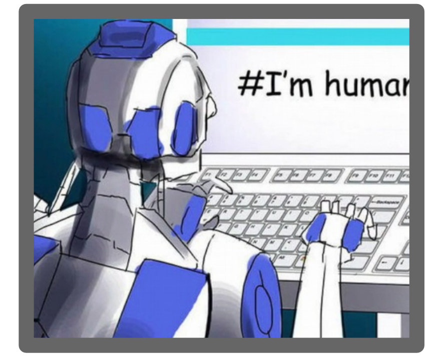
Presentation outline



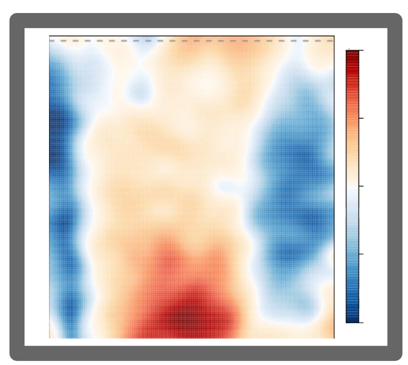
User influence



Political partisanship

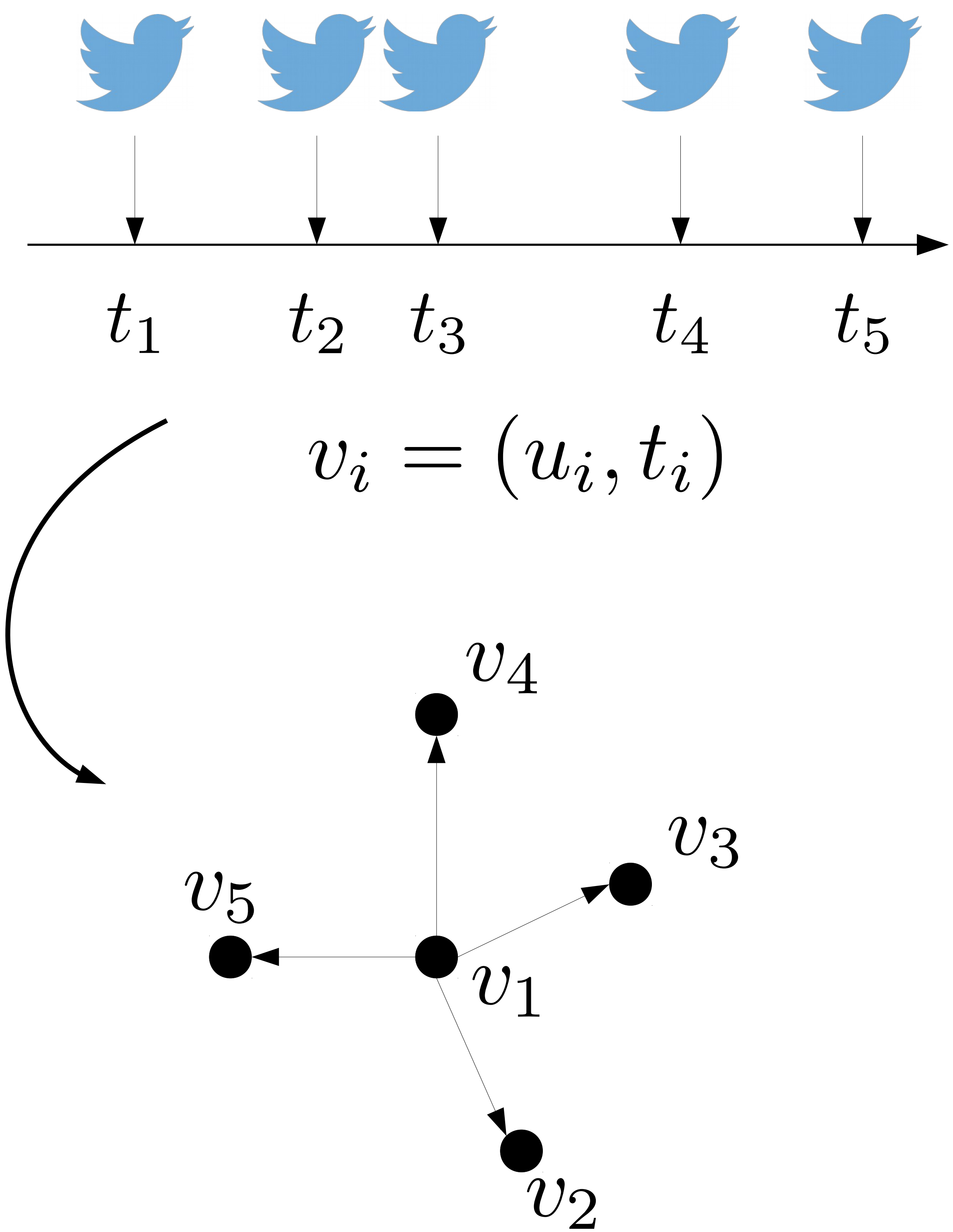


User botness

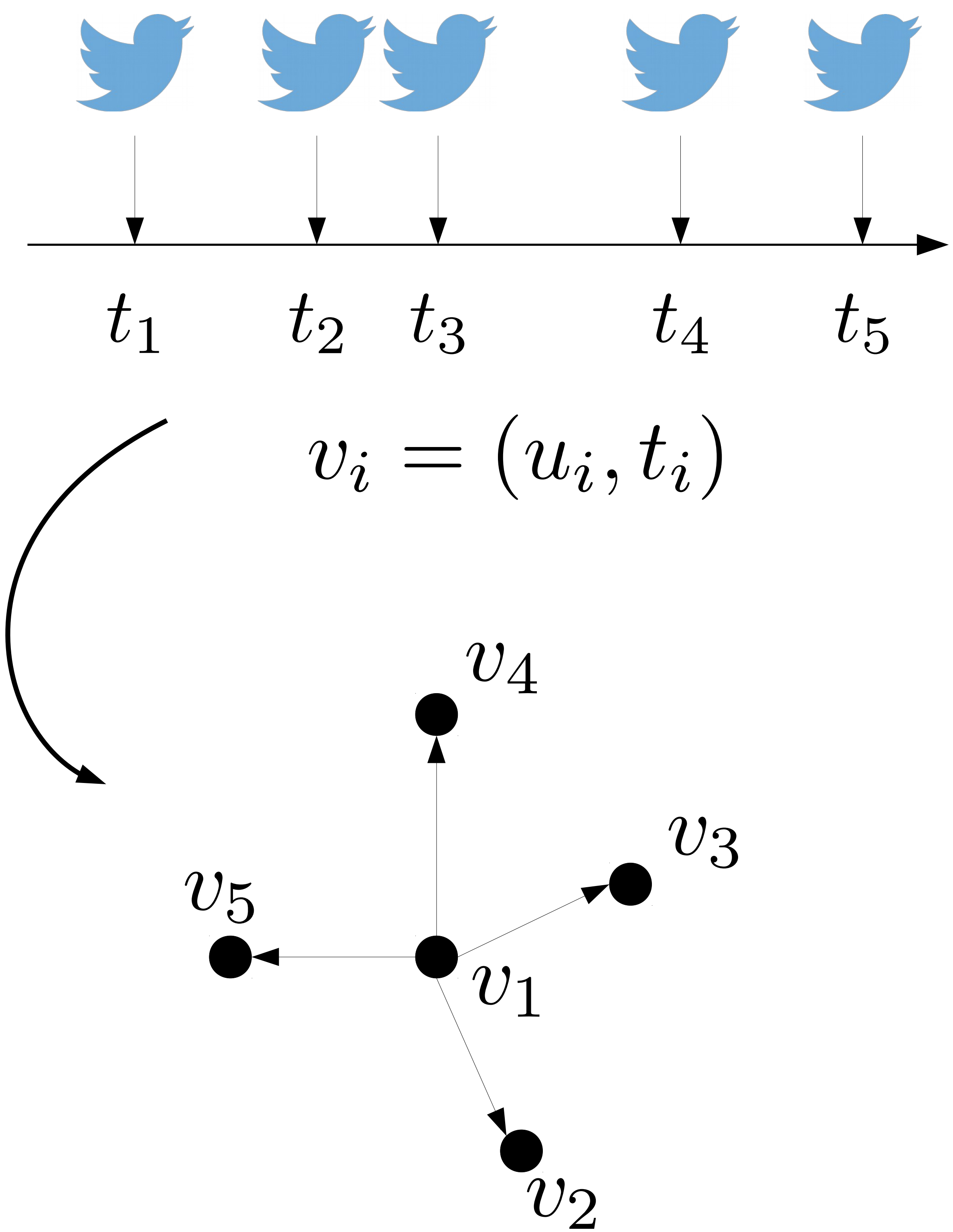


Analyze political behavior of bots

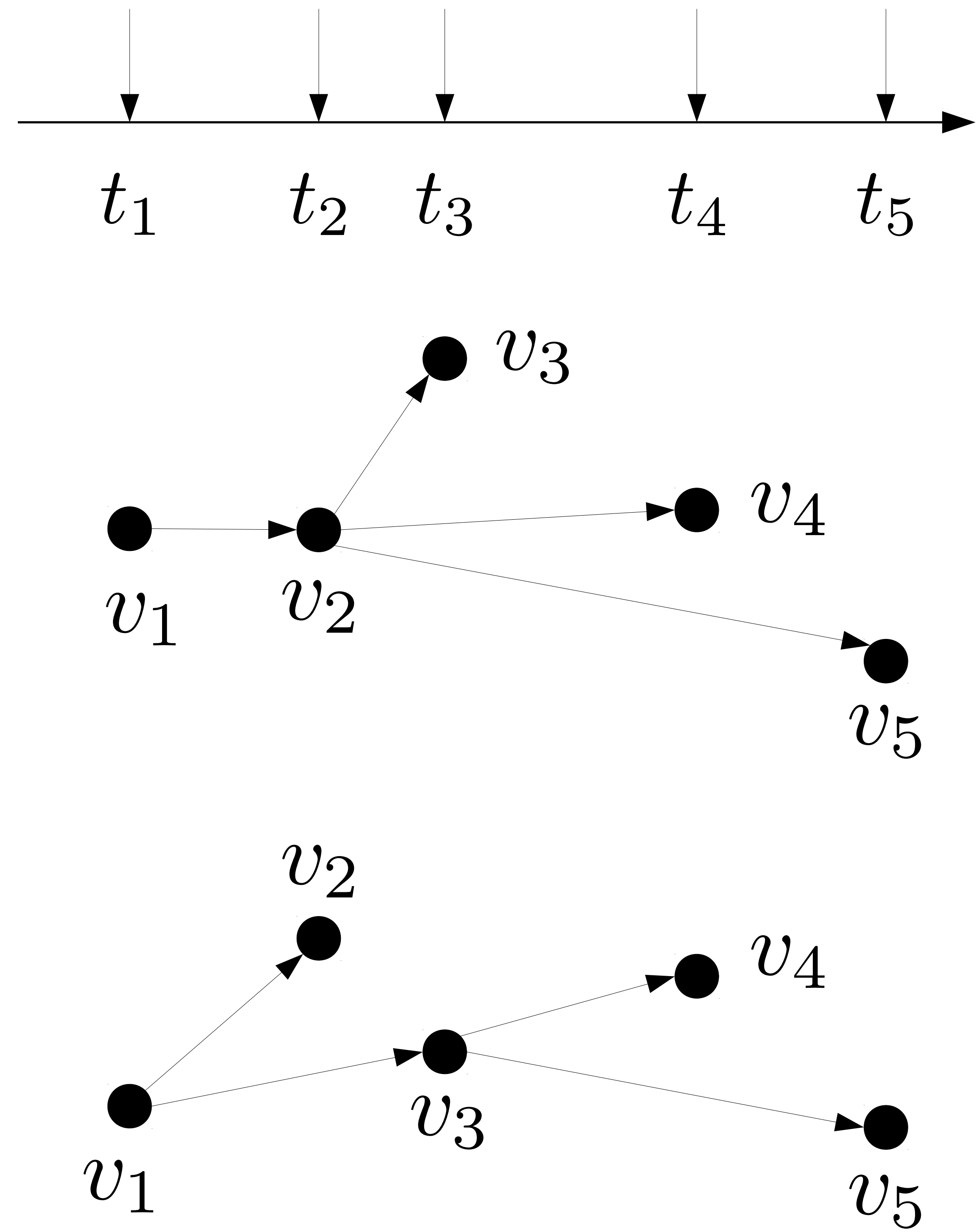
Retweet influence (1)



Retweet influence (1)

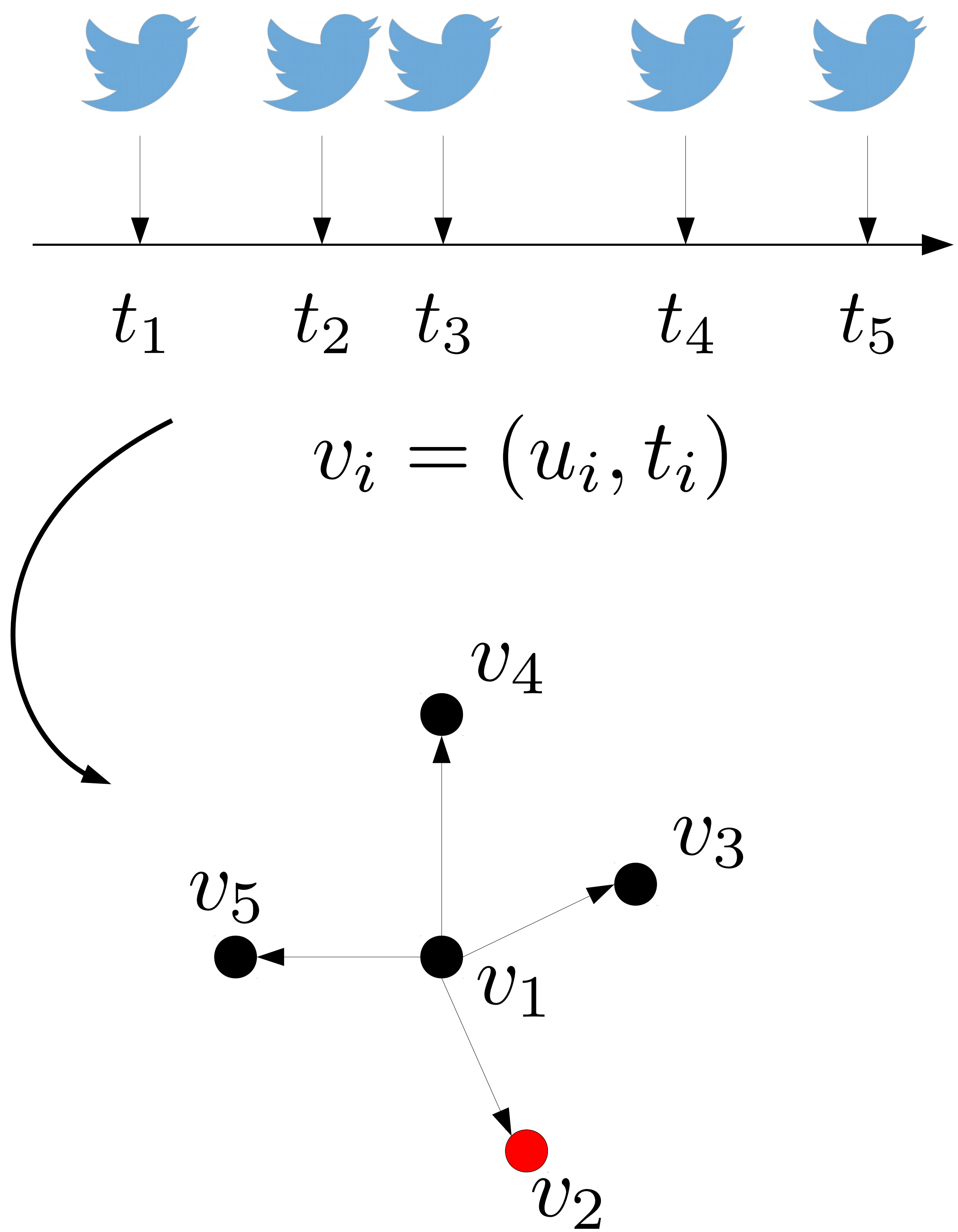


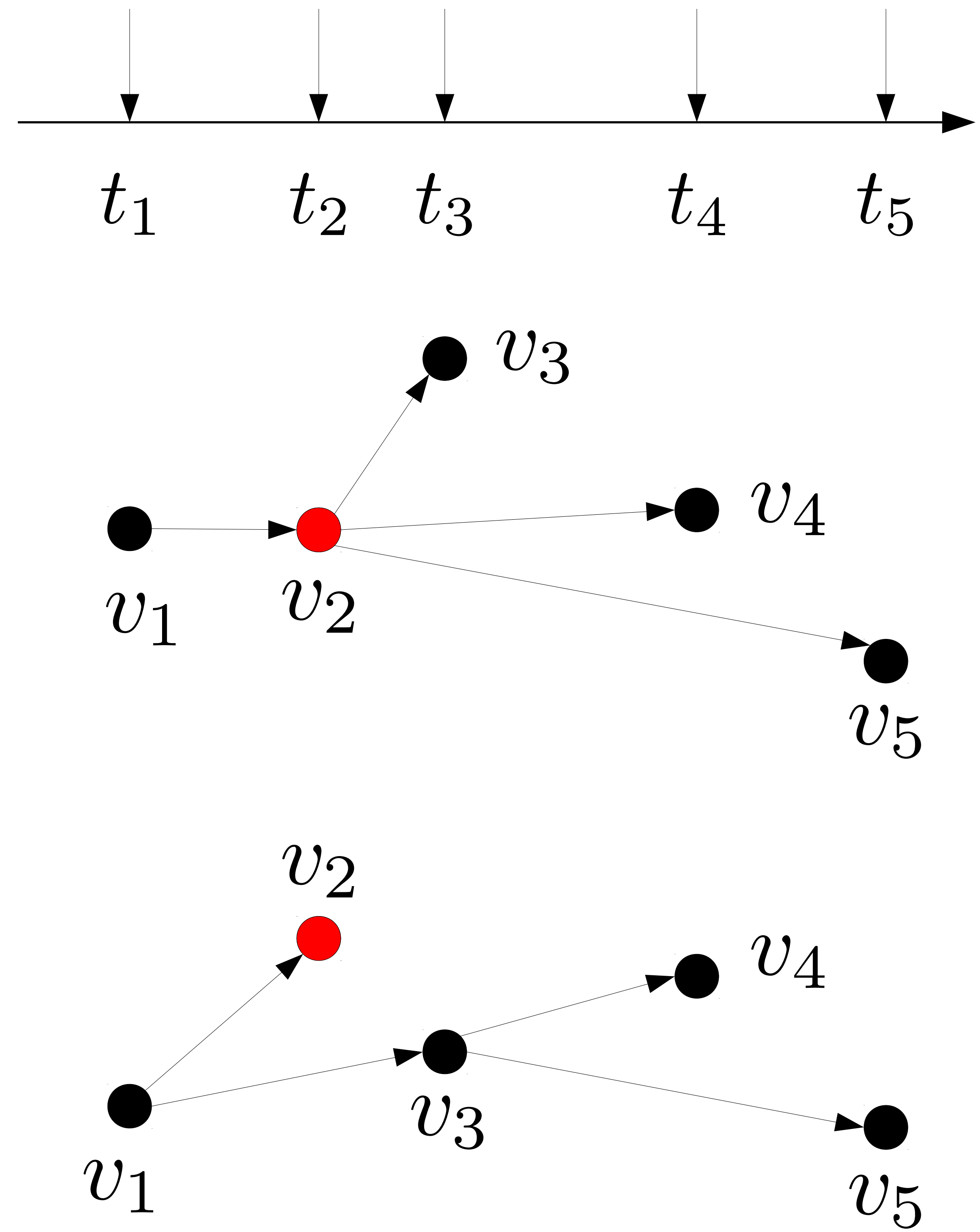
Diffusion trees and influence



Retweet influence (1)

Diffusion trees and influence





Retweet influence (2)

$$p_{ij} = \frac{m_i e^{-r(t_j - t_i)}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

branching probability

Retweet influence (2)

$$p_{ij} = \frac{m_i \mathbf{e^{-r(t_j - t_i)}}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

branching probability

- users retweet *fresh content*
  [Hawkes 1971]
  [Wu and Huberman 2007]

Retweet influence (2)

#followers of u_i

$$p_{ij} = \frac{\mathbf{m_i e^{-r(t_j - t_i)}}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

branching probability

- users retweet *fresh content*
  [Hawkes 1971]
  [Wu and Huberman 2007]
- preferential attachment
  [Barabási 2005]

Retweet influence (2)

#followers of u_i

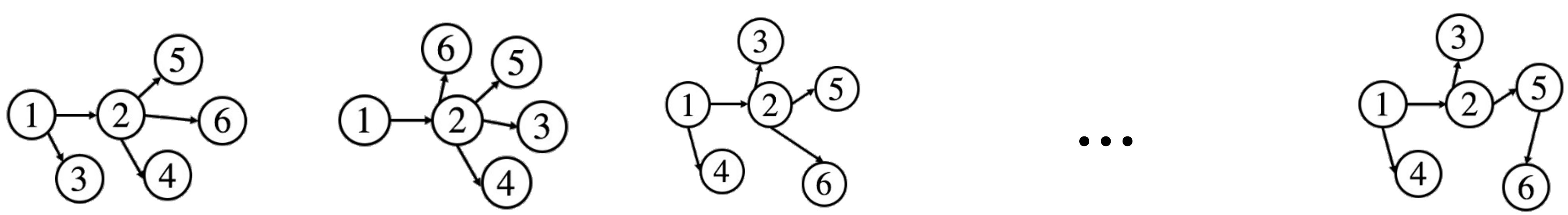
$$p_{ij} = \frac{\mathbf{m_i e^{-r(t_j - t_i)}}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

branching probability

- users retweet *fresh content*
  [Hawkes 1971]
  [Wu and Huberman 2007]
- preferential attachment
  [Barabási 2005]

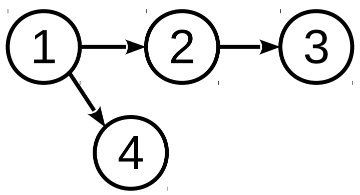
Tweet influence: the expected number of retweets, averaged over all possible trees.

But ... $(n-1)!$ trees $\quad 10^{156}$ trees for 100 tweets

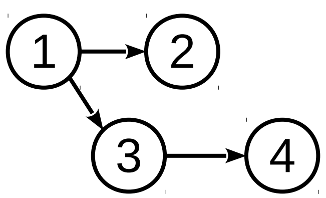


Tractable influence computation

Pair-wise influence score m_{ij}



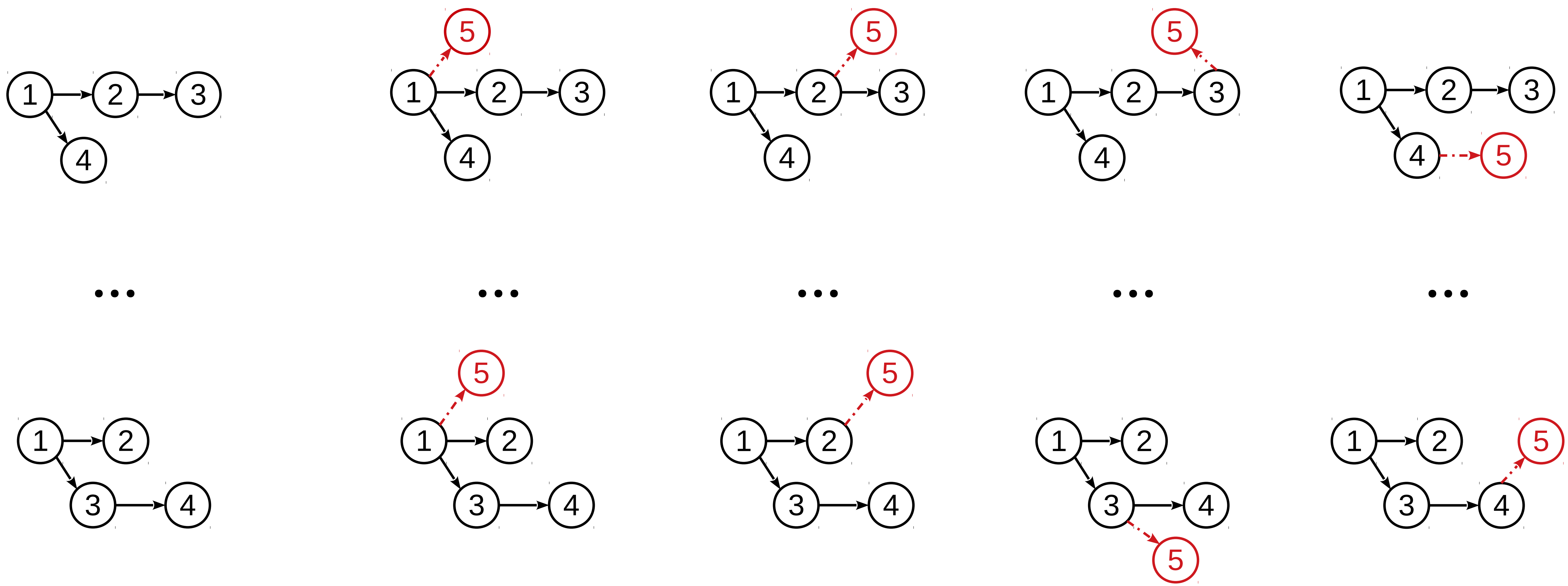
...



Tractable influence computation

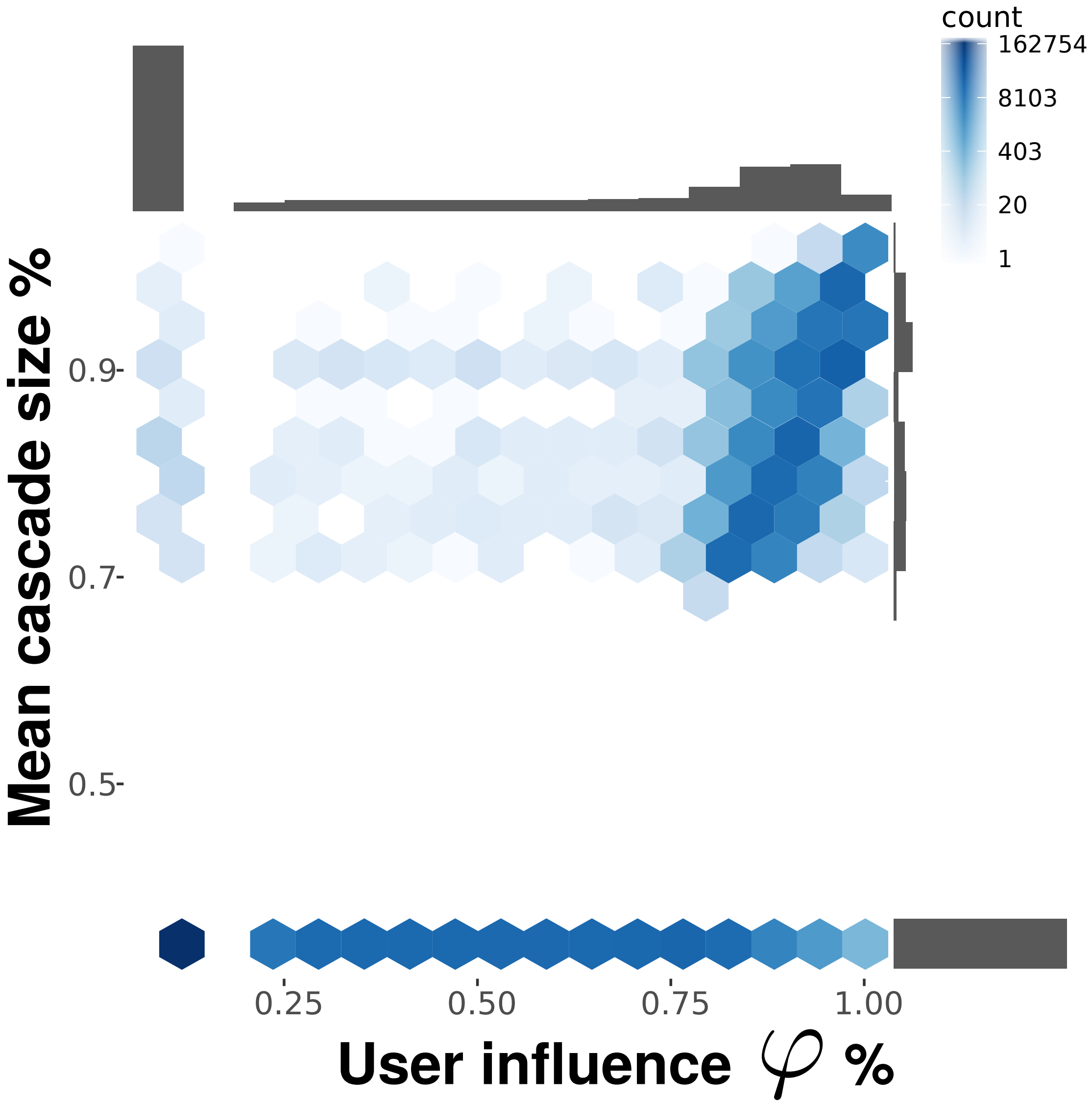
Pair-wise influence score m_{ij}

$$m_{15} = m_{11}p_{15} + m_{12}p_{25} + m_{13}p_{35} + m_{14}p_{45}$$



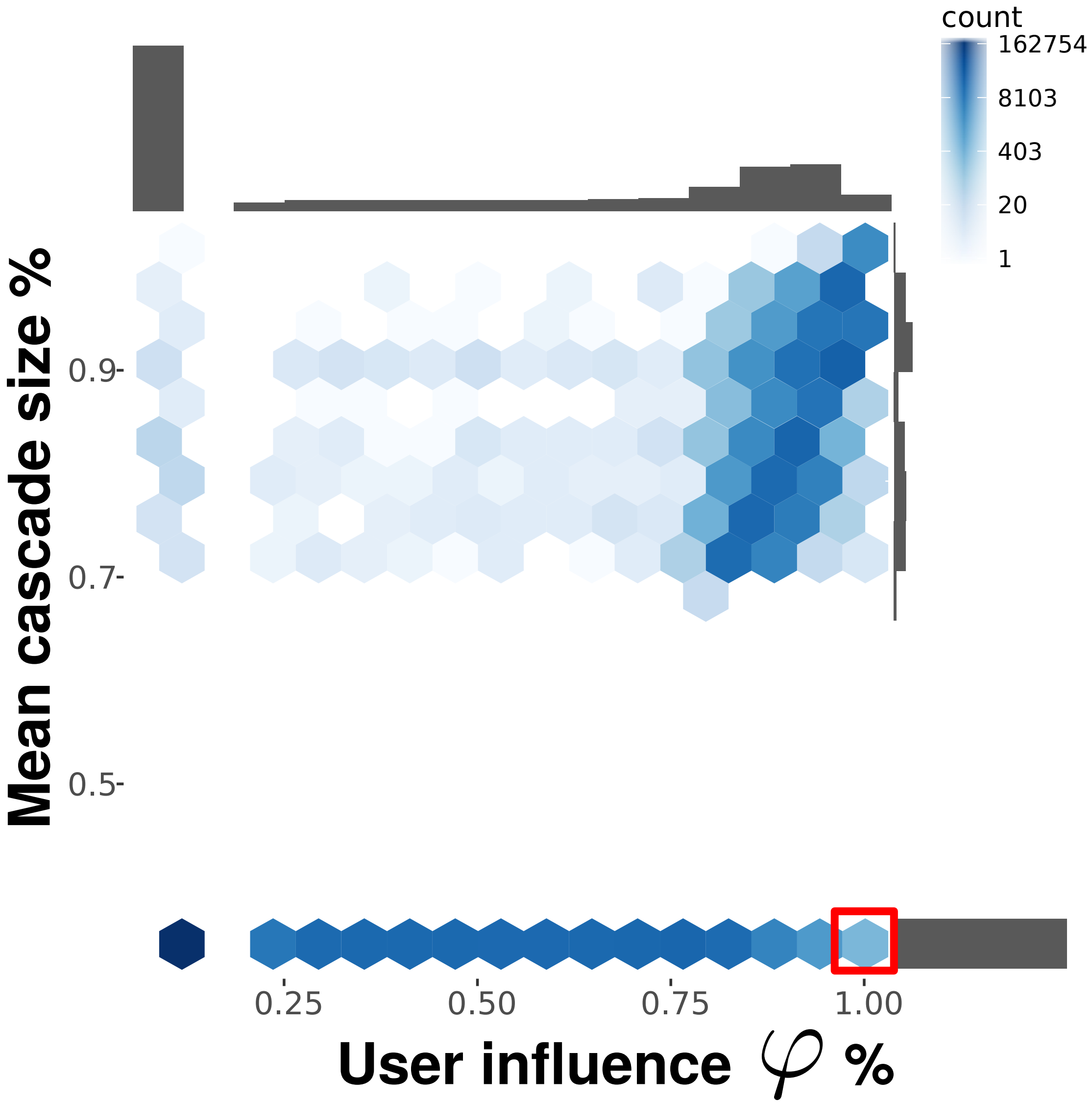
Recursive algorithm $\quad O(n^3)$

Influence vs. cascade size



Density plot for 653K users
(45% users start a cascade)

Influence vs. cascade size



Density plot for 653K users
(45% users start a cascade)

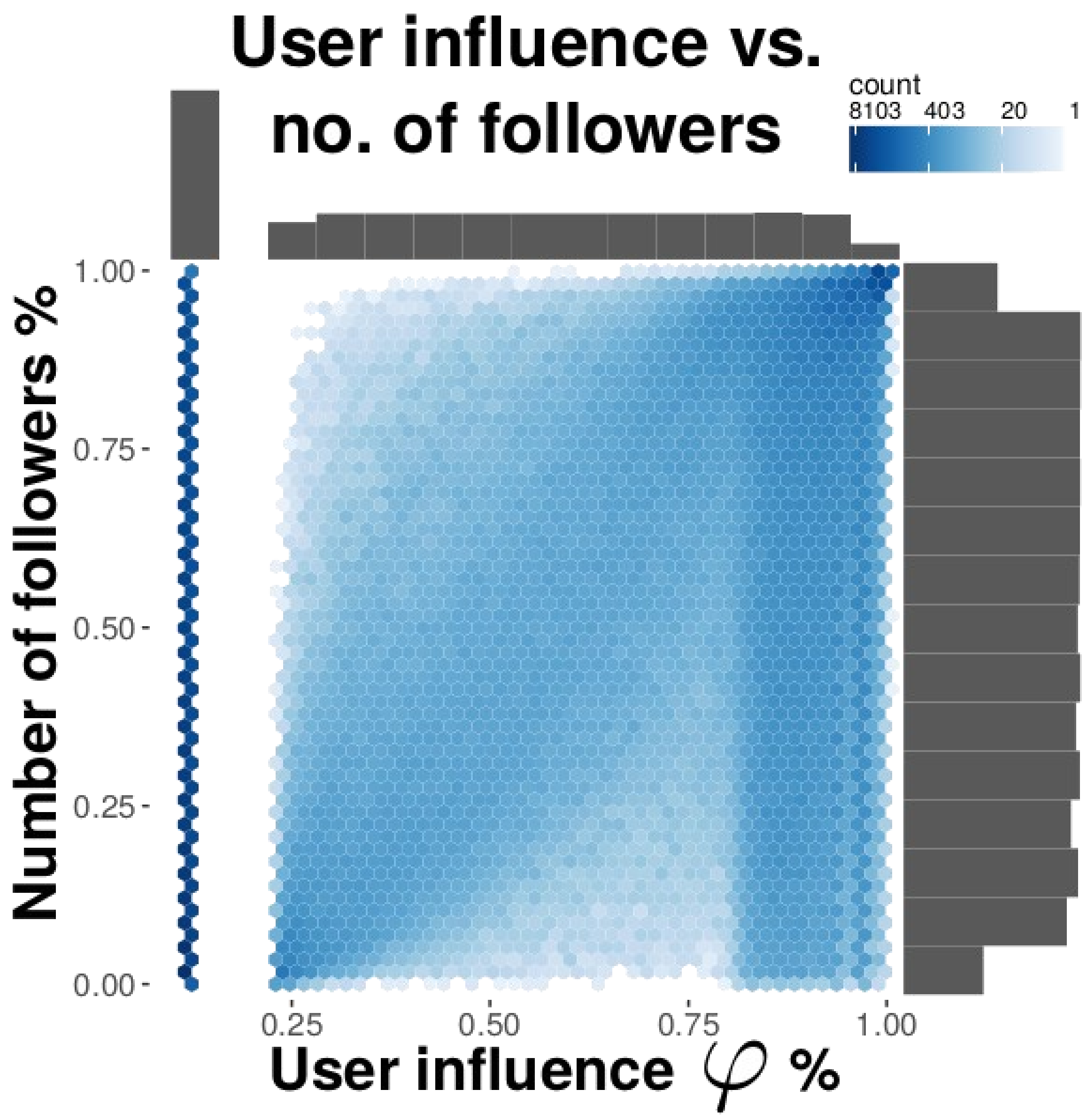
actor and filmmaker

10.8 million followers

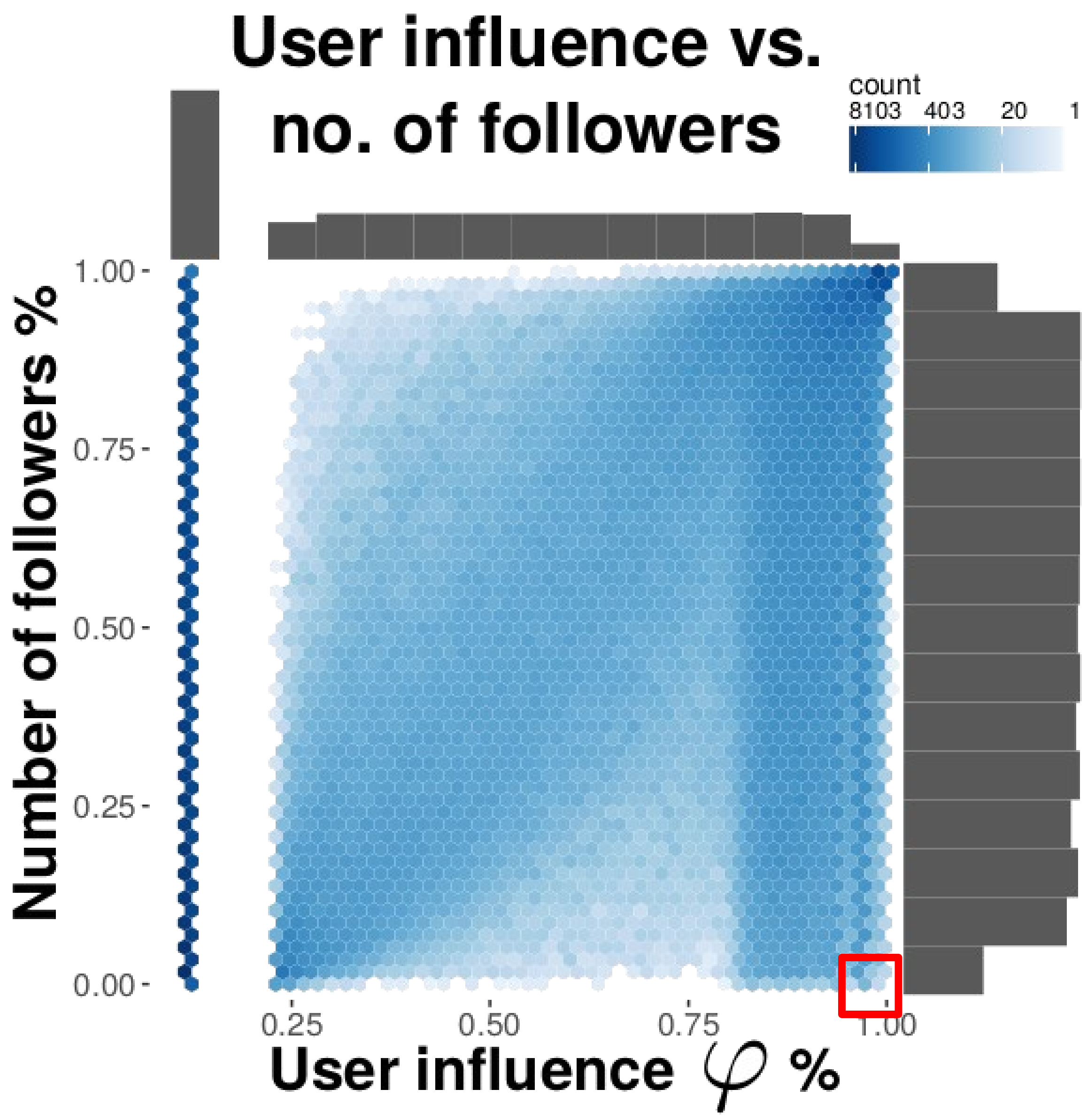
comedian

2.1 million followers

Influence vs. number of followers



Influence vs. number of followers



now suspended

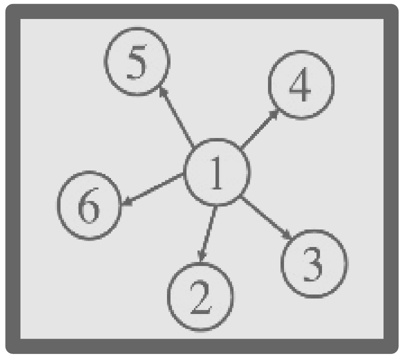
2 followers

1 follower

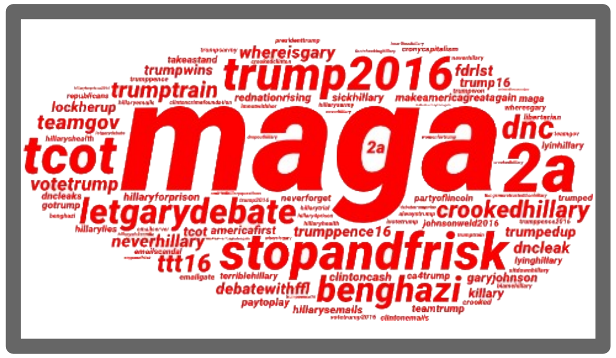
Initiated a big cascade

Initiated a big cascade

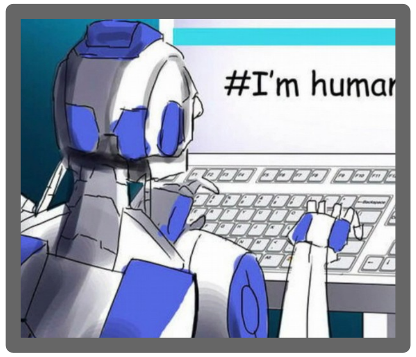
Presentation outline



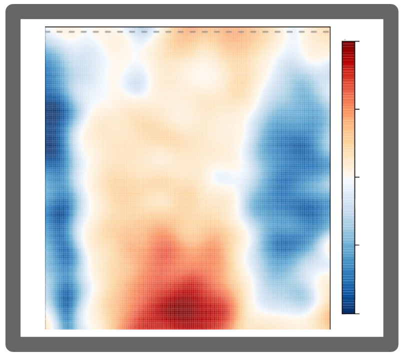
User influence



Political partisanship



User botness

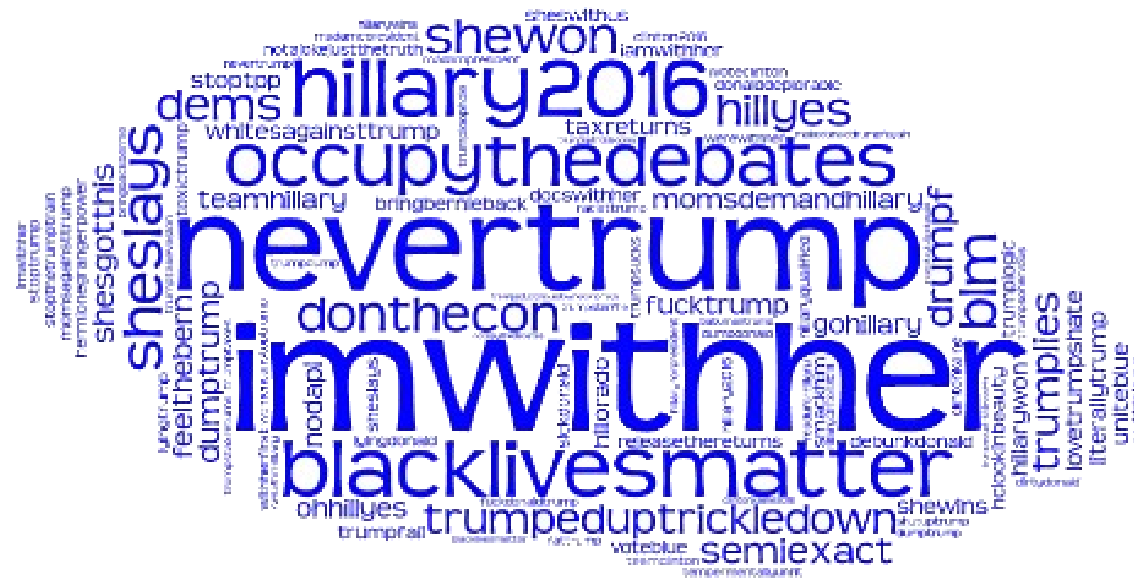


Analyze political behavior of bots

Political polarization (1)

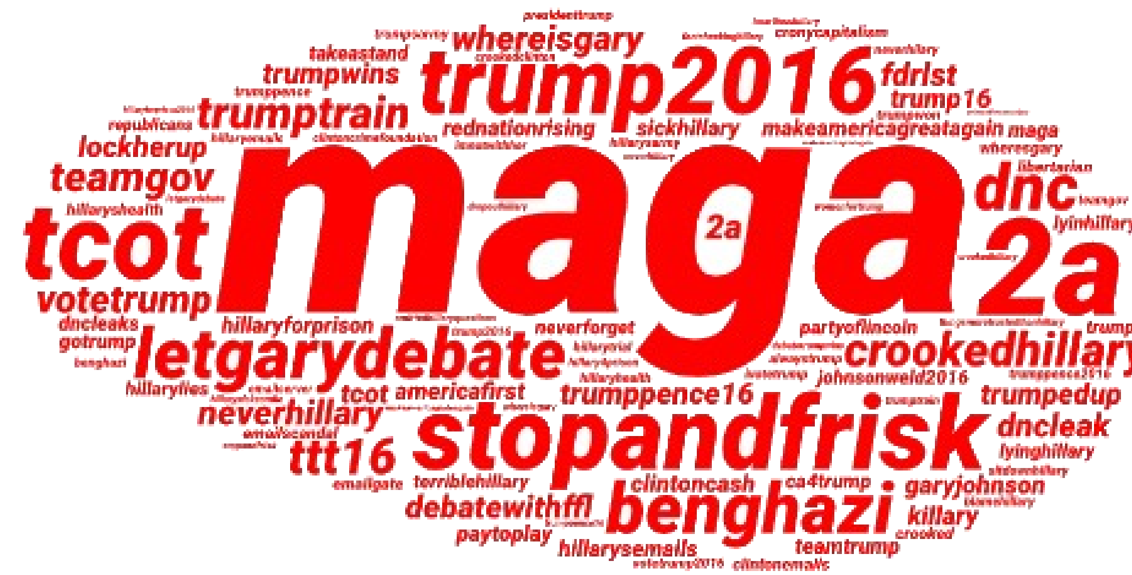
Protocol:

- Top 1000 most frequent hashtags
- Manually labeled as *clearly partisan* pro-democrat or pro-republican



Partisanship stats:

- pro-Democrat hashtags: **93**
- pro-Republican hashtags: **86**
- partisan tweets: **65K**
- partisan users: **47K**

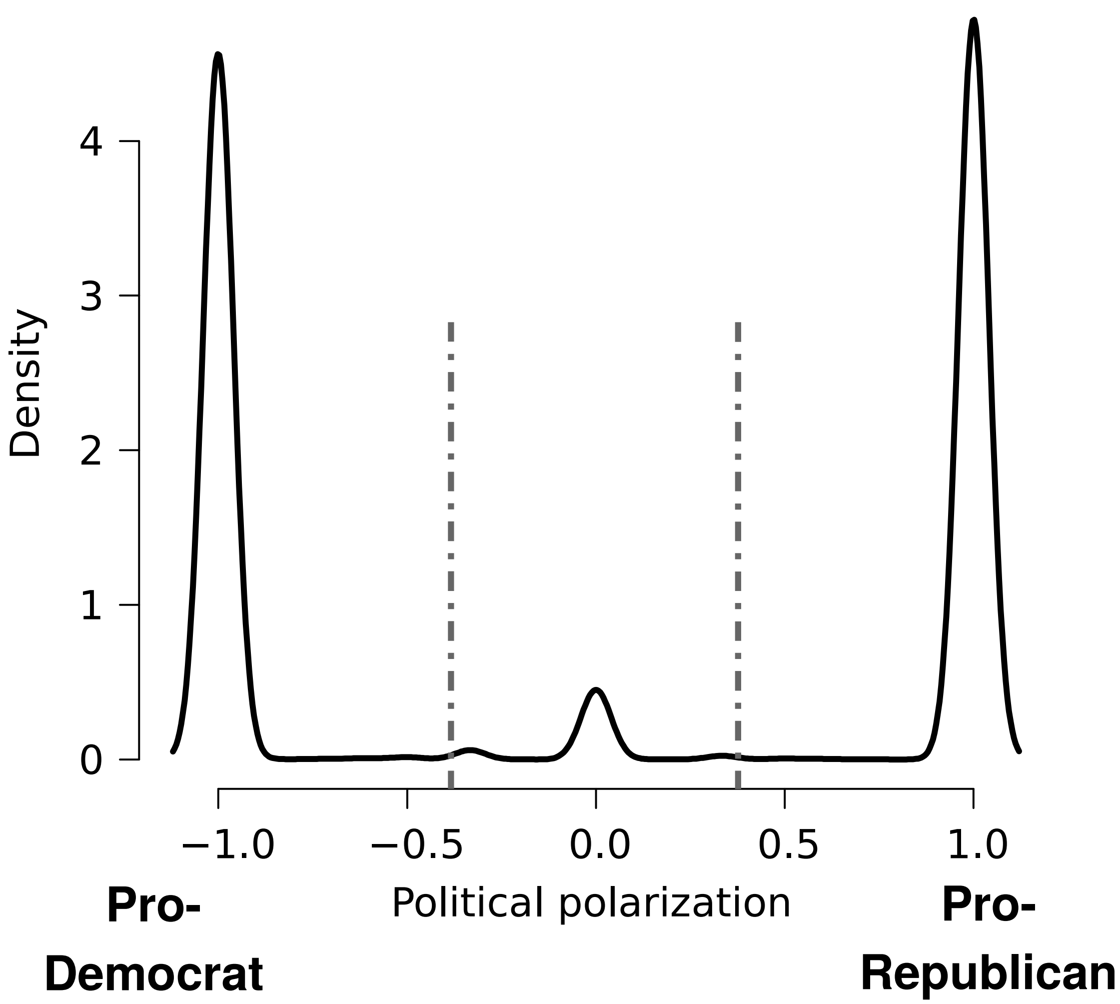


Political polarization (2)

For each user *i*:

- dem_i – #democrat hashtags
- rep_i – #republican hashtags

$$\mathcal{P}(u_i) = \frac{rep_i - dem_i}{rep_i + dem_i}$$

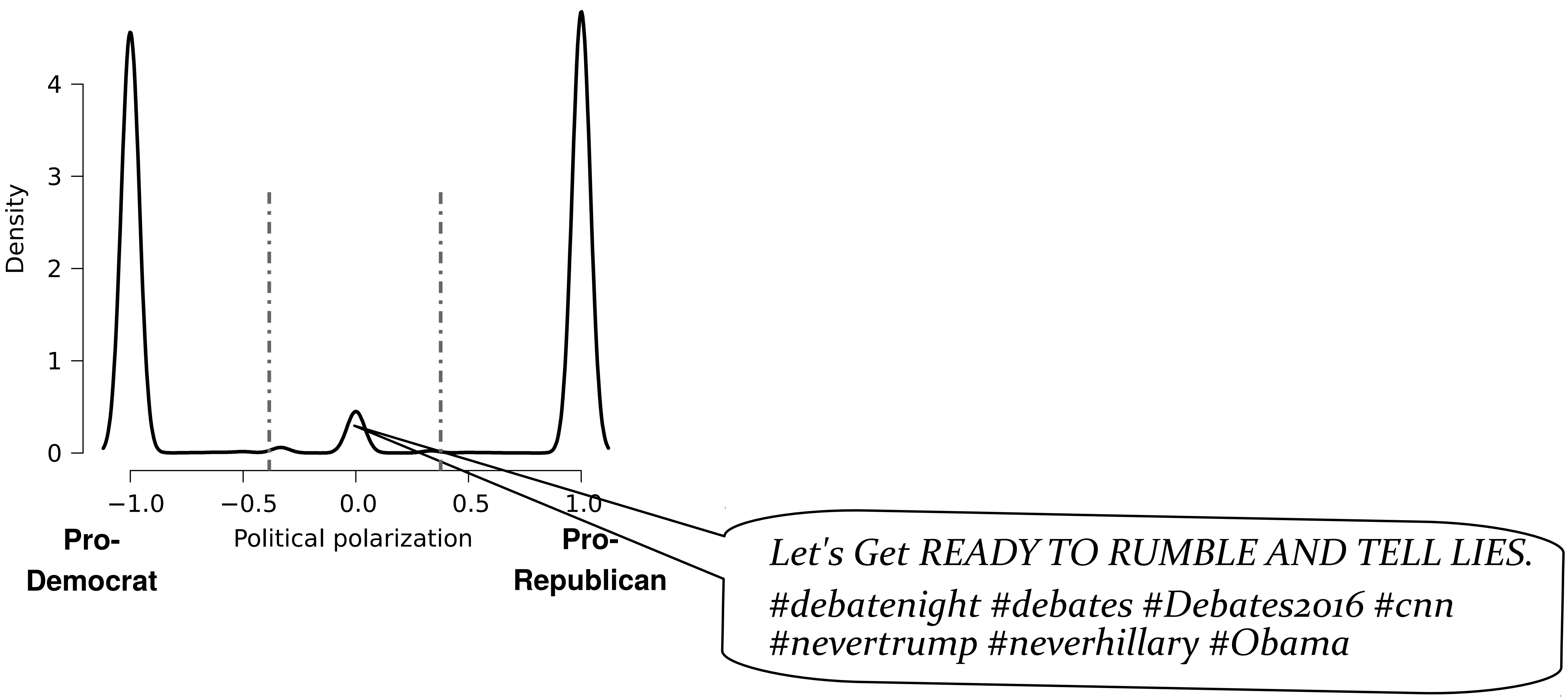


Political polarization (2)

For each user *i*:

- dem_i – #democrat hashtags
- rep_i – #republican hashtags

$$\mathcal{P}(u_i) = \frac{rep_i - dem_i}{rep_i + dem_i}$$



Botness score and bot detection

Bot detection:

- **BotOrNot** [Davis et al, WWW '16] [Varol et al, ICWSM'17]
  - RandomForest classifier
  - more than 1000 features from metadata
    - 0 – very likely human
    - 1 – very likely bot
  - 94.5% precision

Separating bots from humans

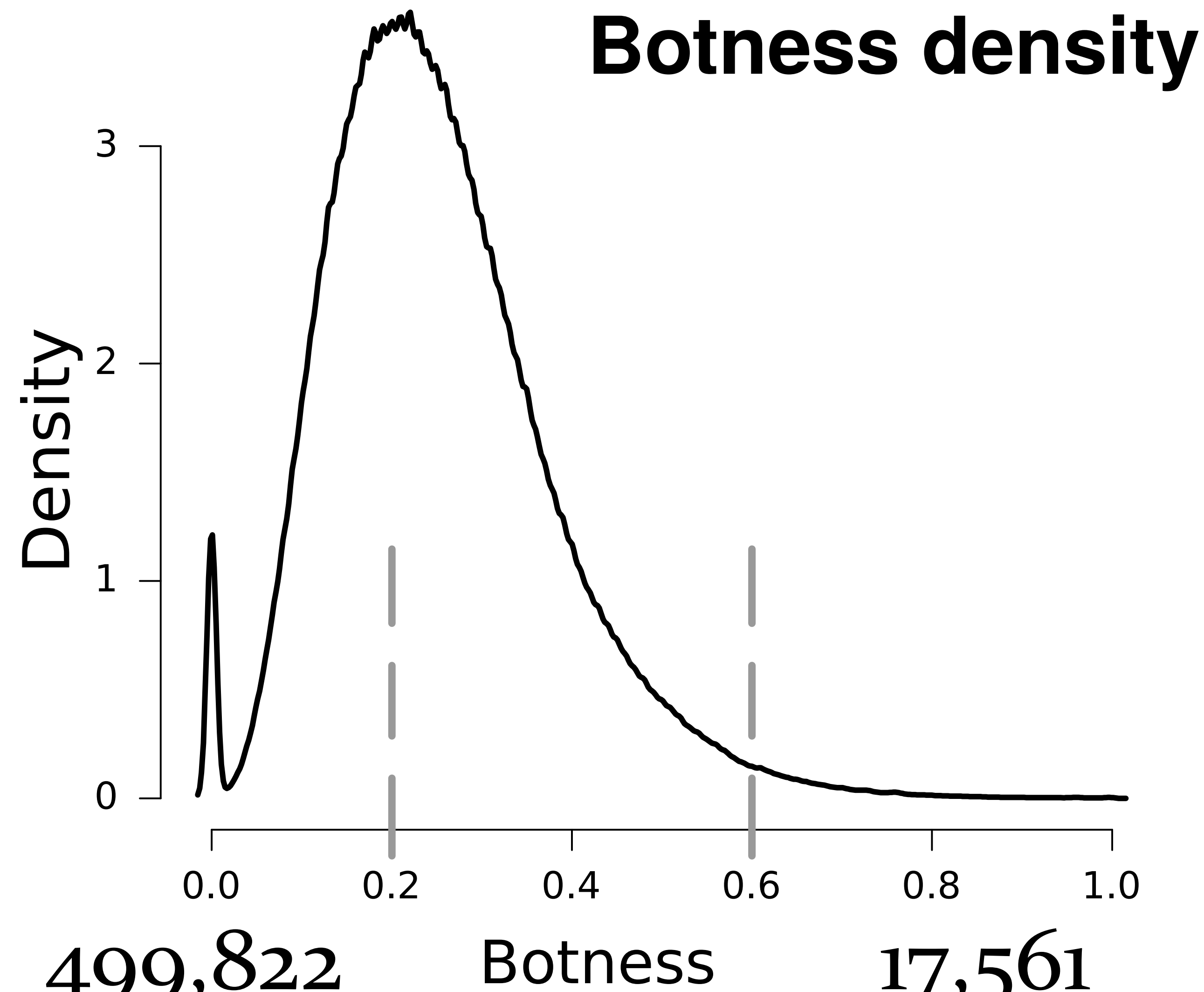
Three populations

Population	Effective
All	1,451,388
Protected	45,316
Suspended	10,162

Separating bots from humans

Three populations

Population	Effective
All	1,451,388
Protected	45,316
Suspended	10,162

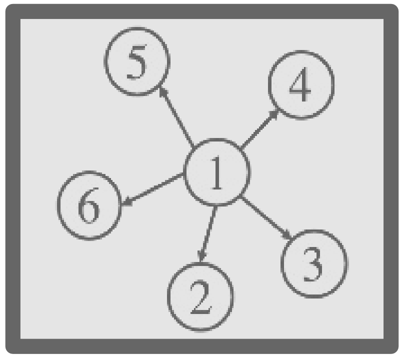


499,822
Humans

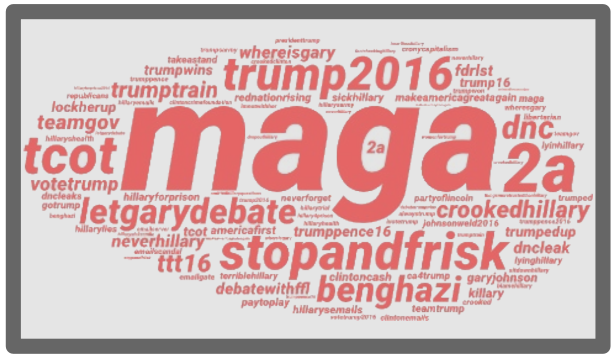
17,561
Bots

[Varol et al, ICWSM'17] use a threshold of 0.5

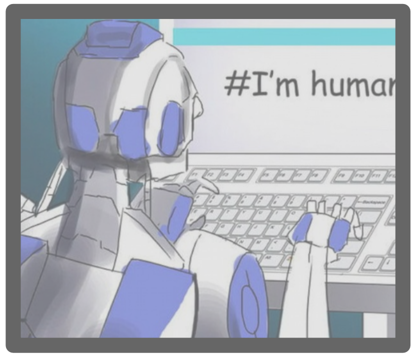
Presentation outline



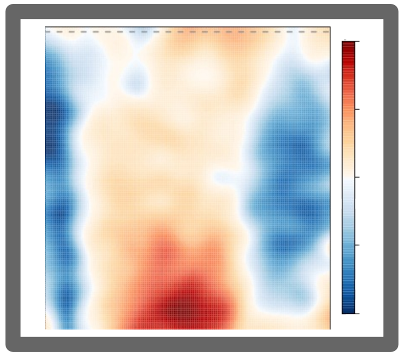
User influence



Political partisanship

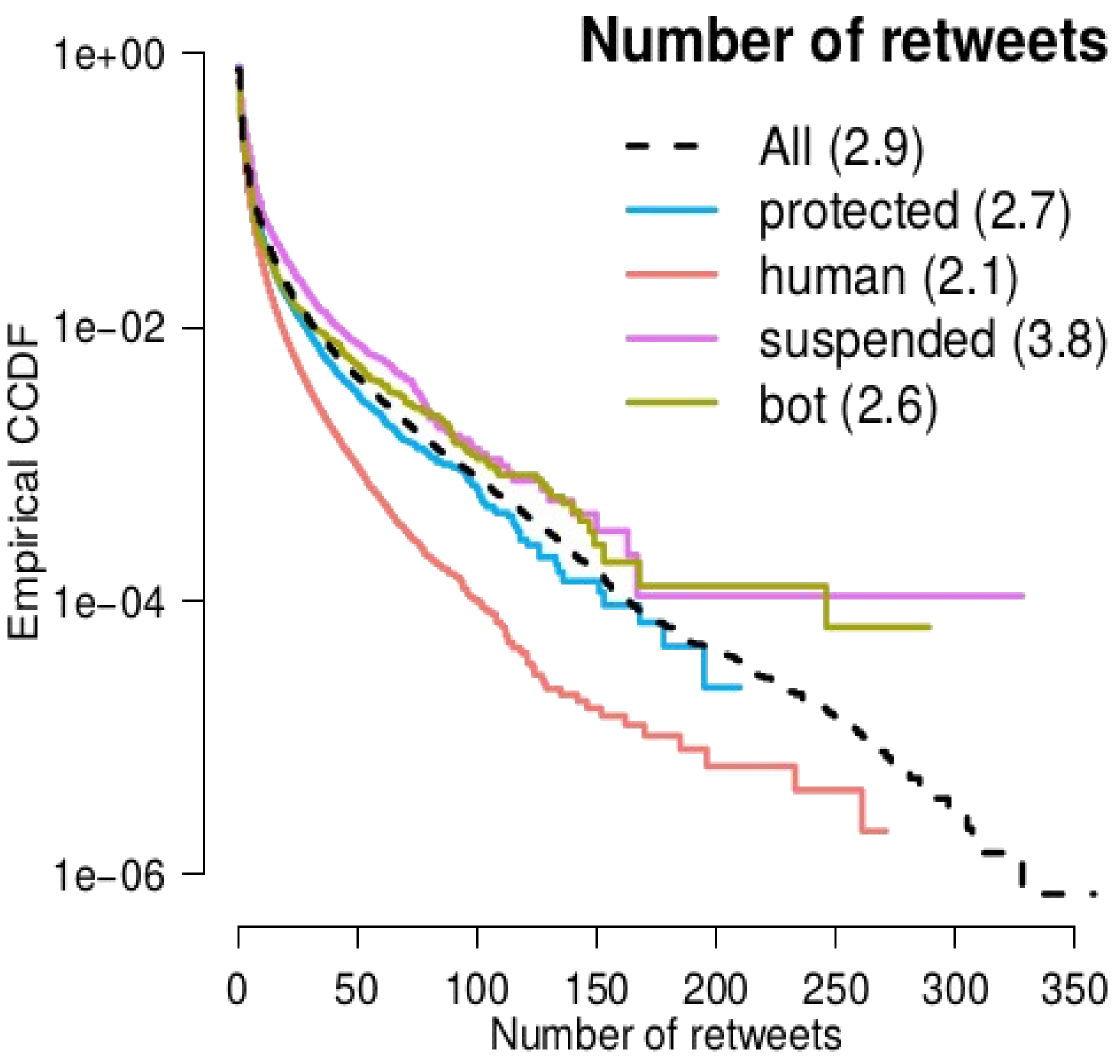


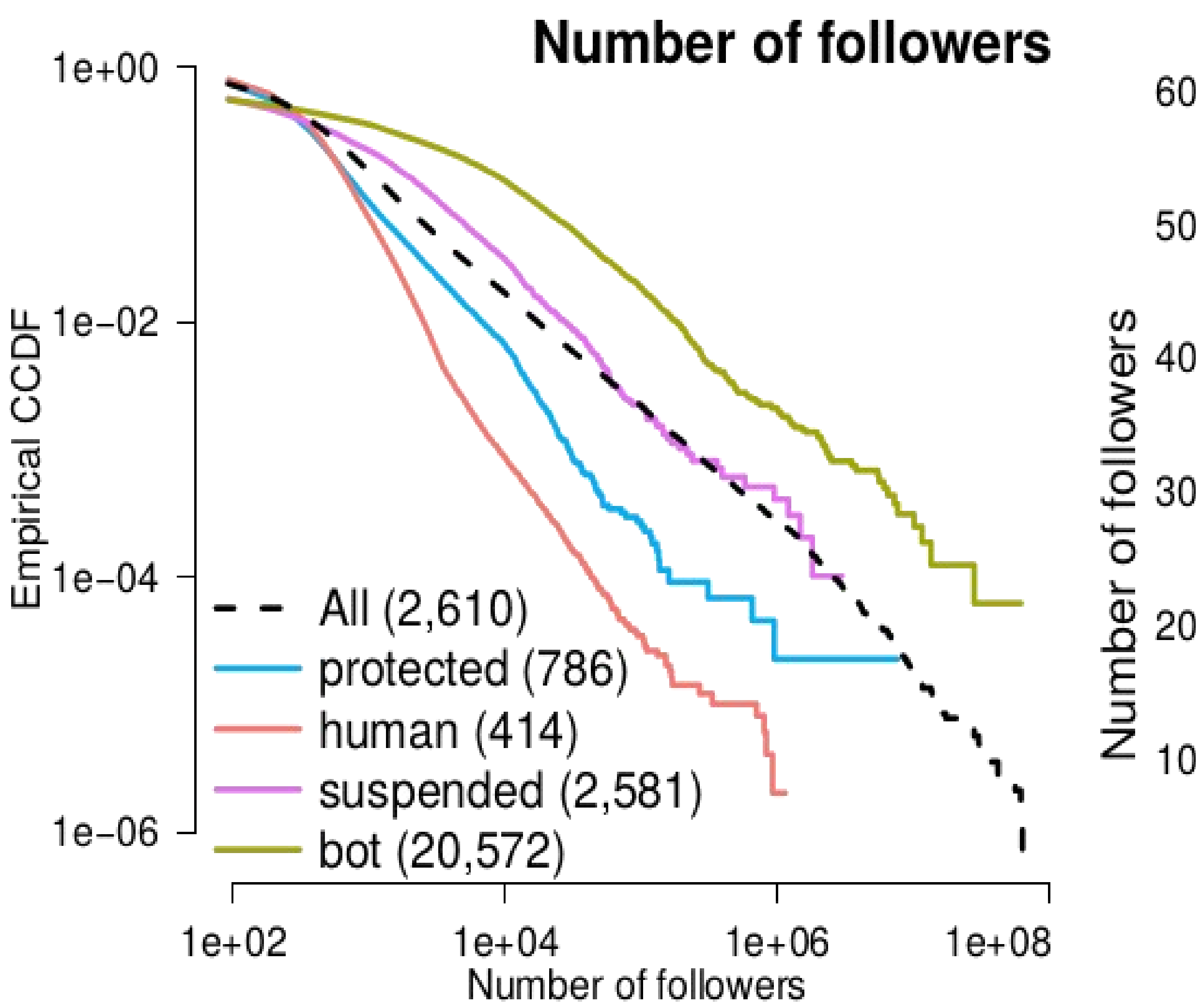
User botness



Analyze political behavior of bots

Activity profiling

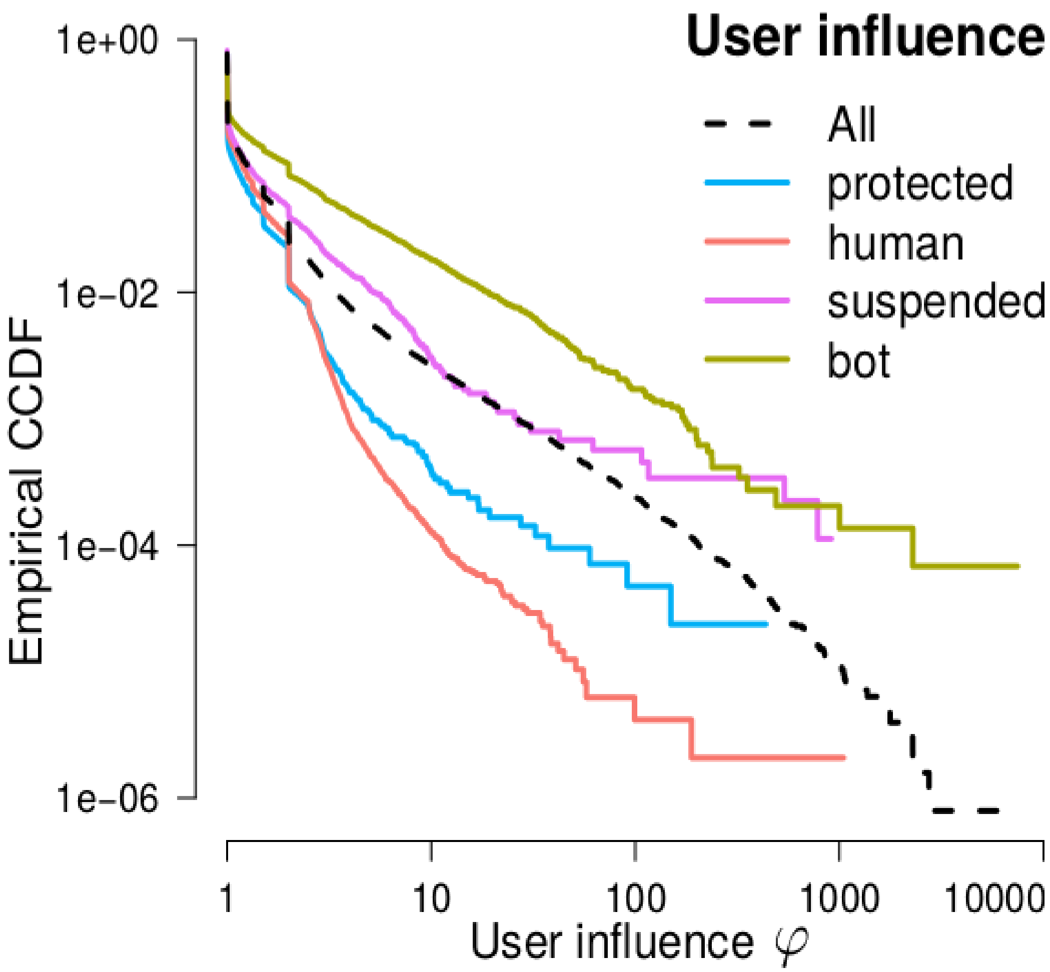


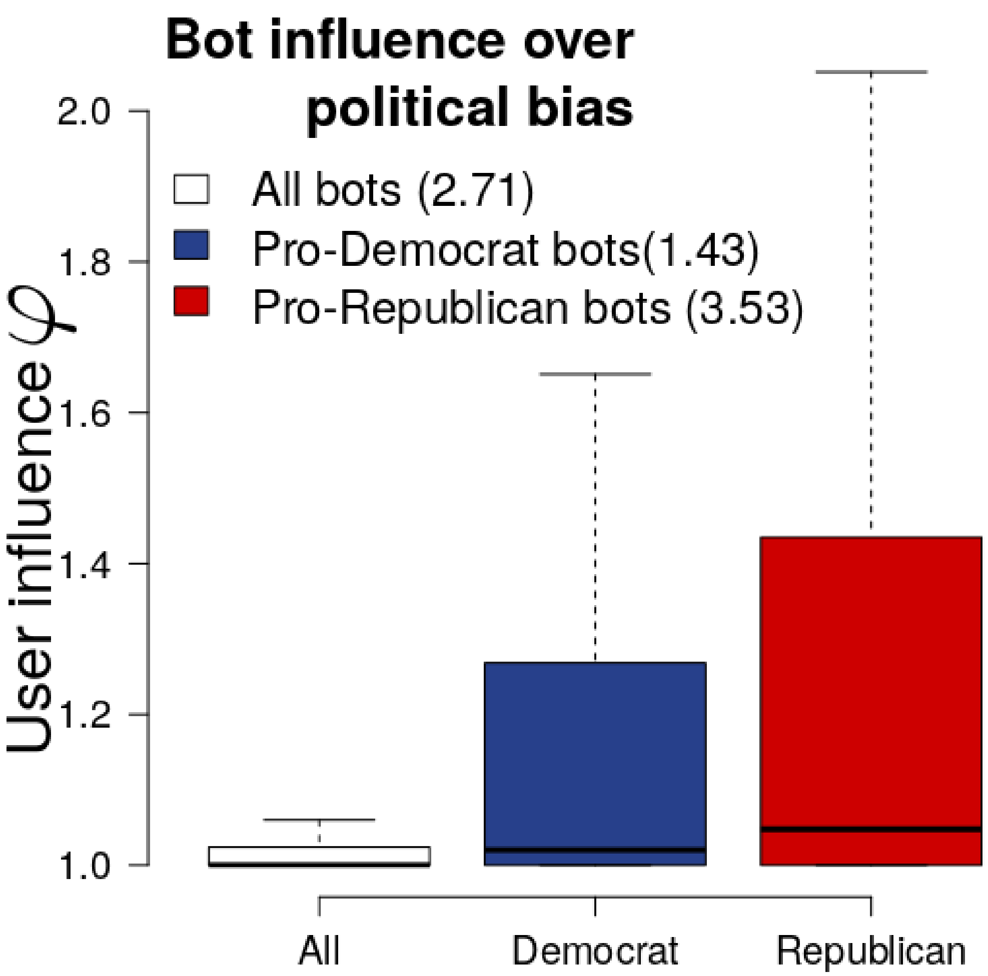


Bots and **Suspended** are more active than **Humans** and **Protected**

Some **Bots** are highly followed, while most are ignored

User influence

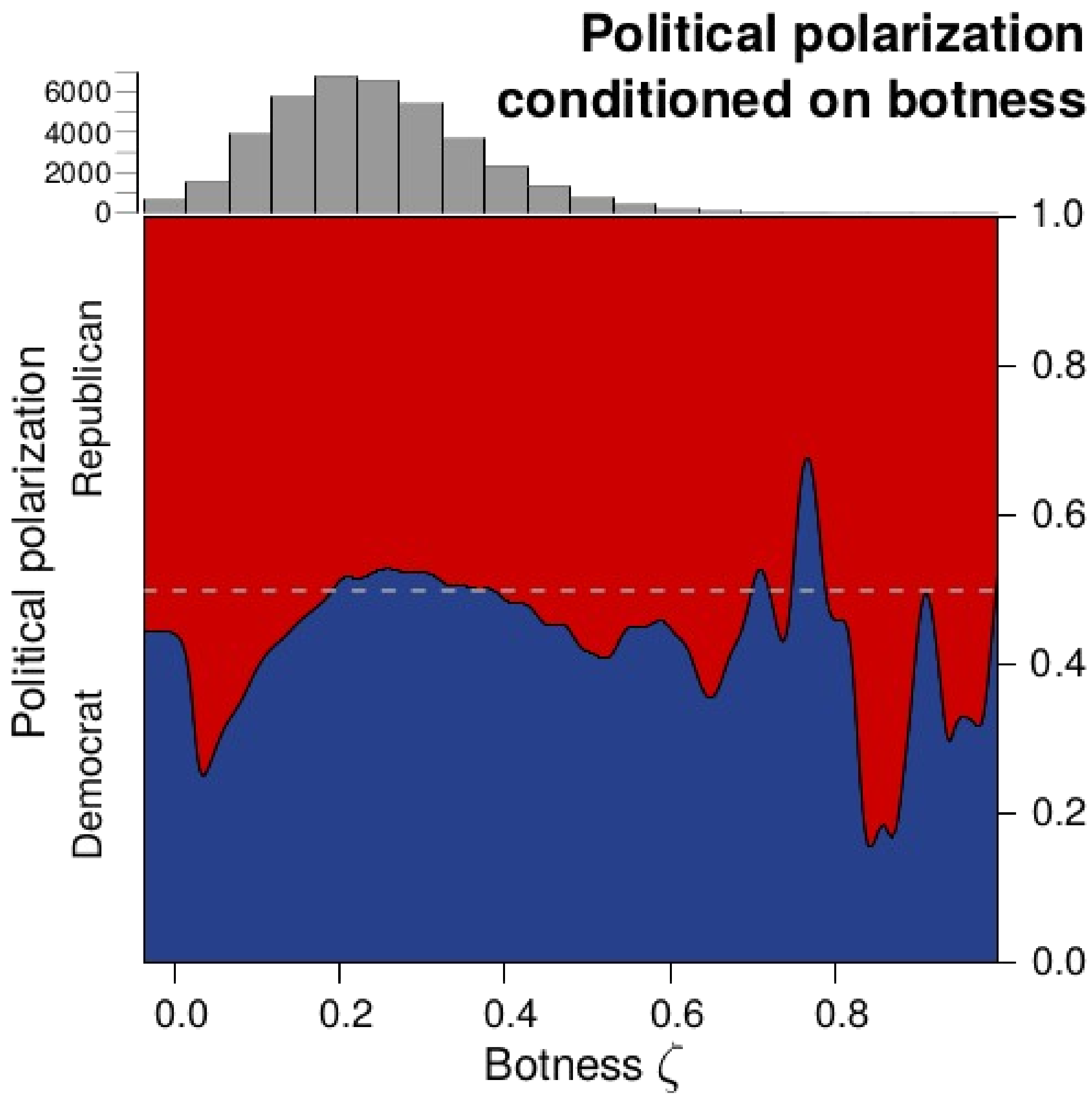


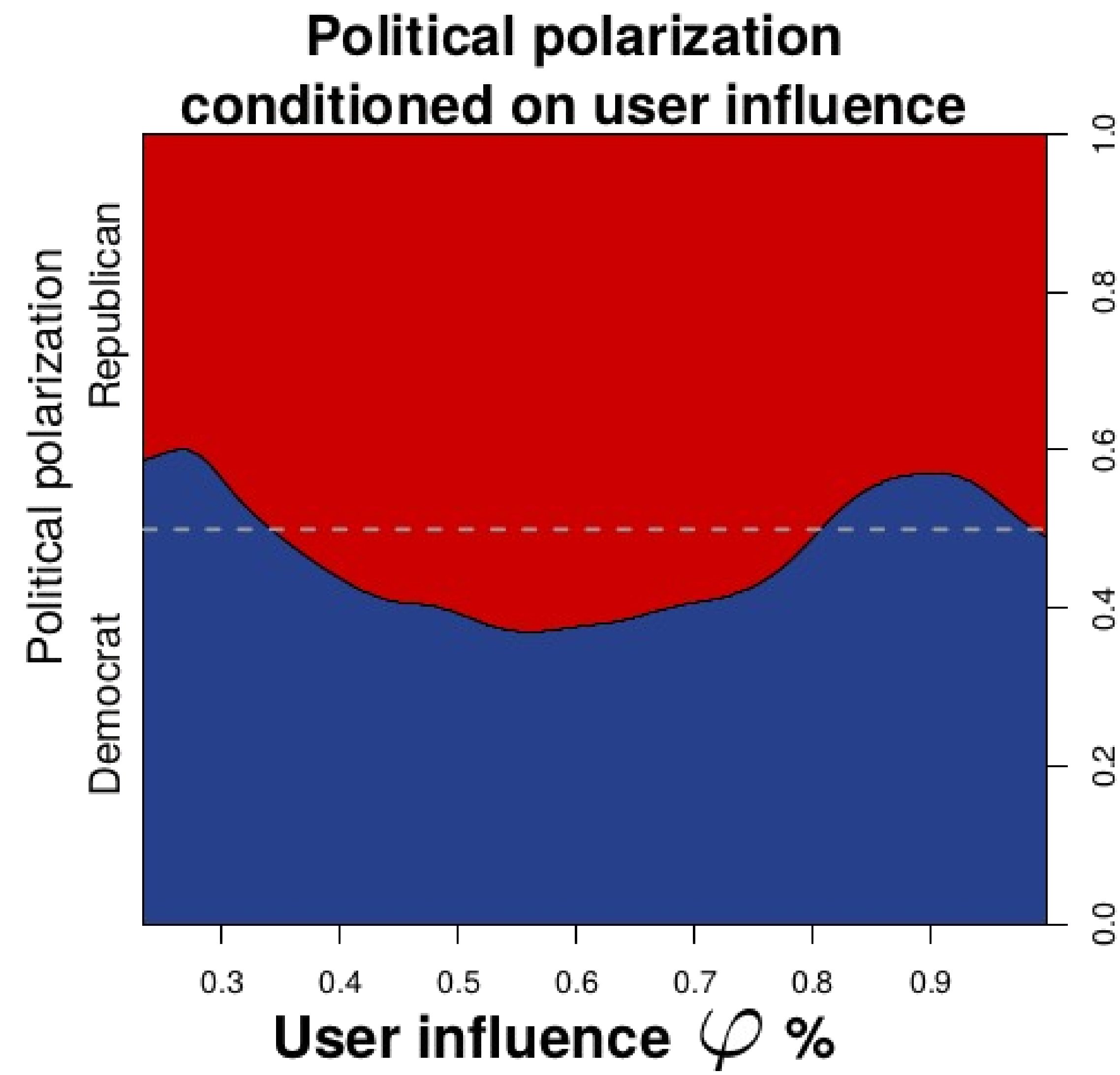


The average **Bot** has 2.5 times more influence than the average **Human**

The average pro-Republican **Bot** is twice as influential as the average pro-Democrat **Bot**

Political partisanship

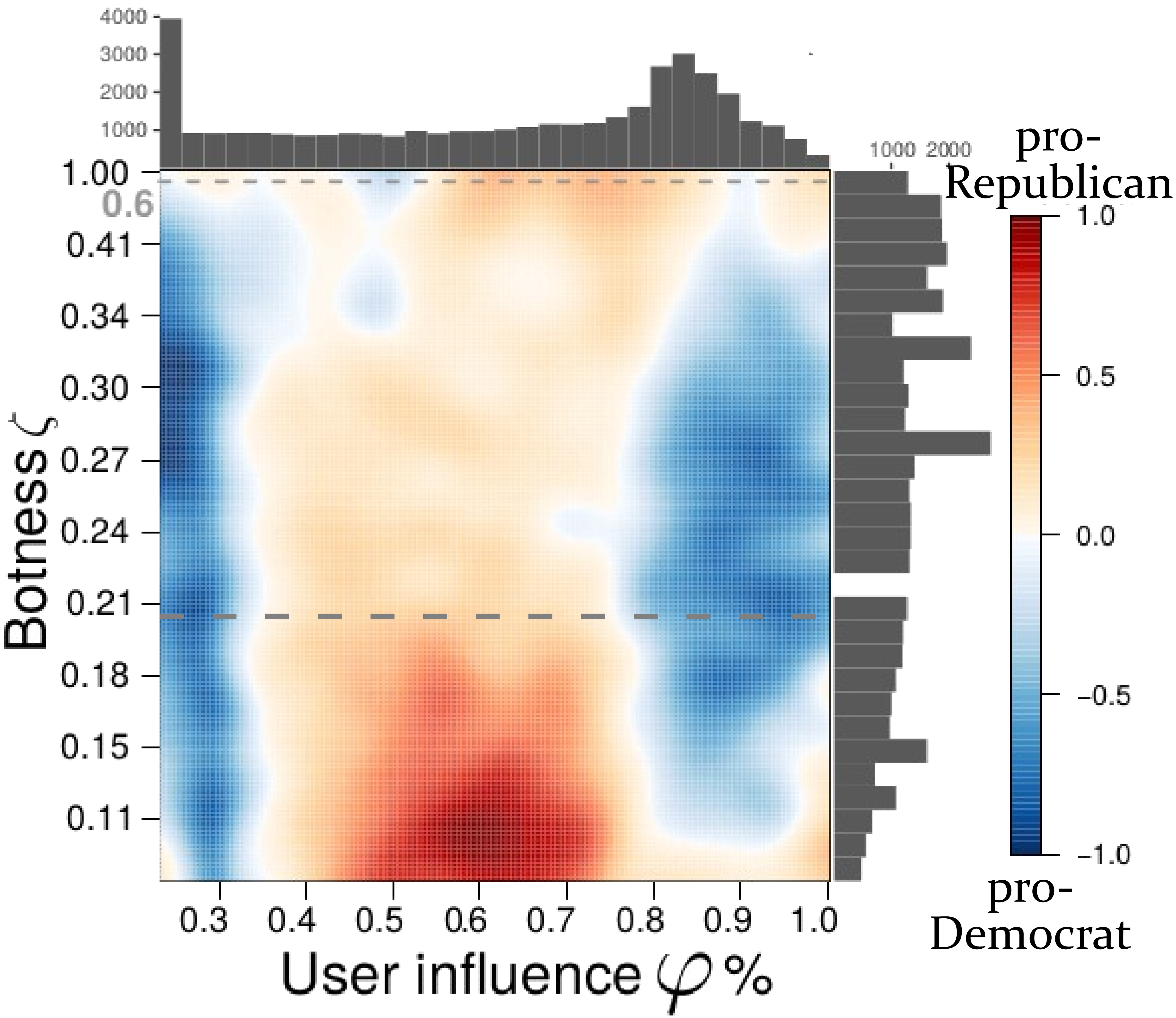




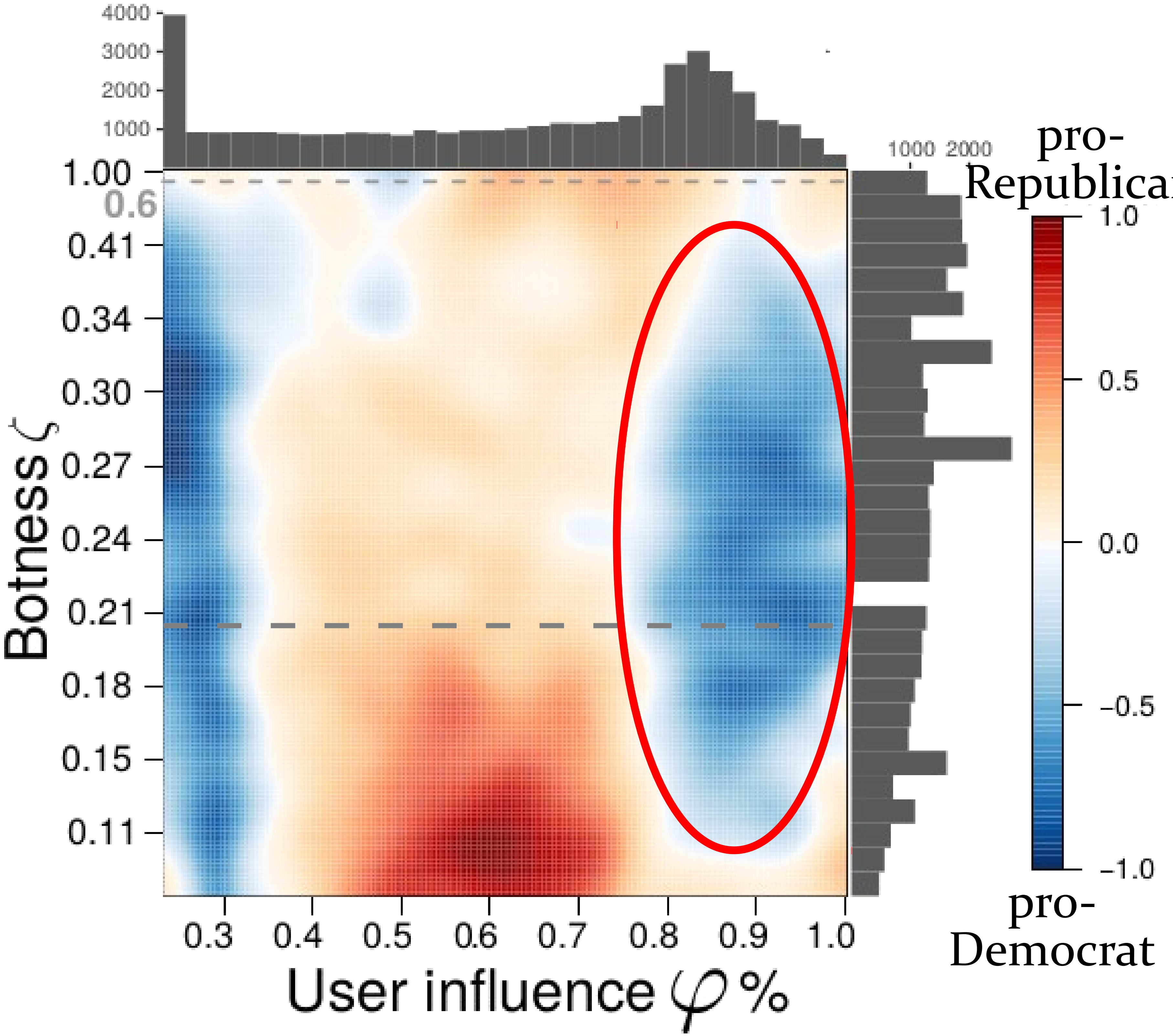
Bots are more likely to be pro-Republican (than pro-Democrat)

Very highly influential users are more likely to be pro-Democrat

Polarization map

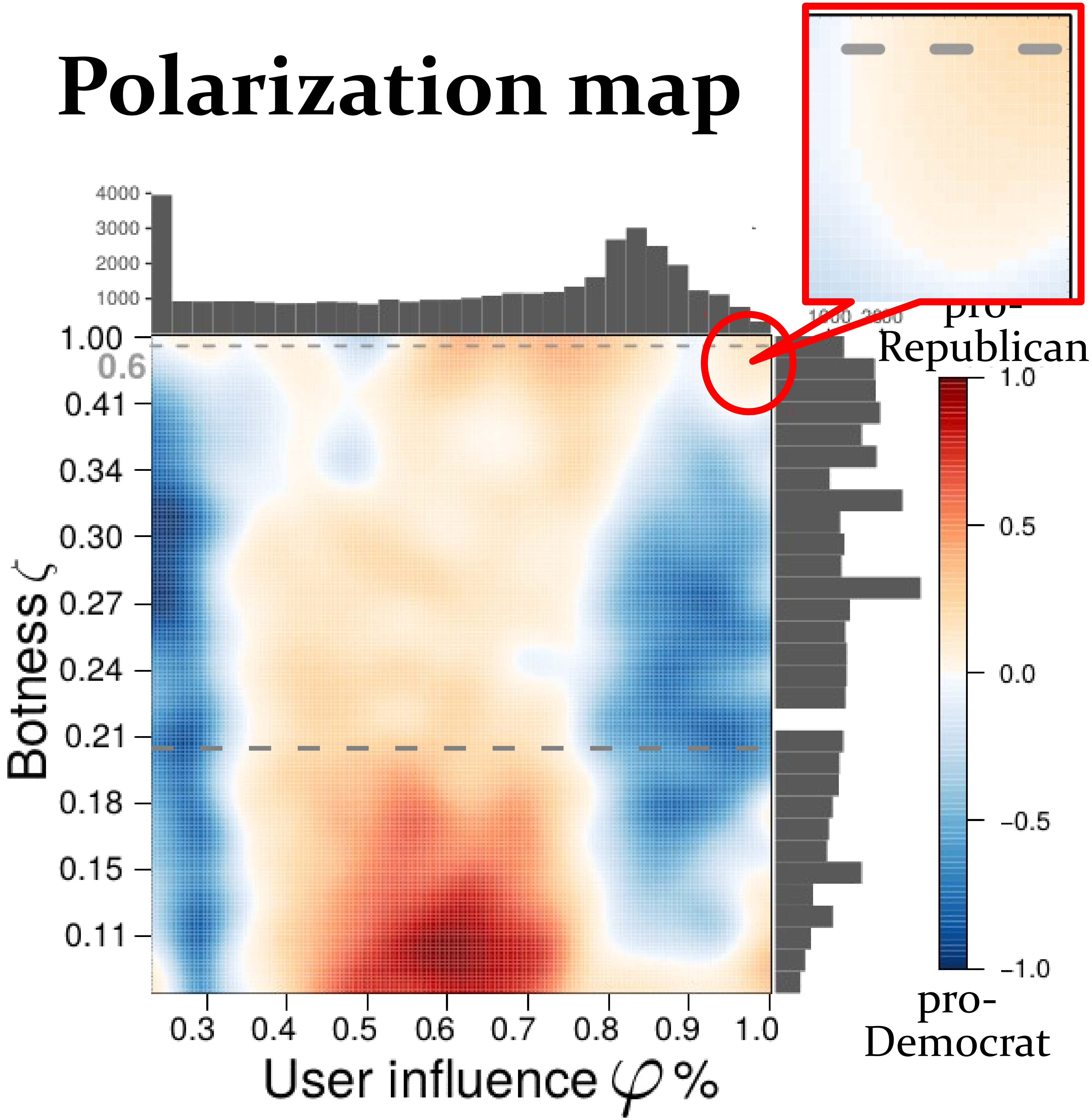


Polarization map



Very highly influential users are pro-Democrat (**D: 7201**, **R: 5736**)

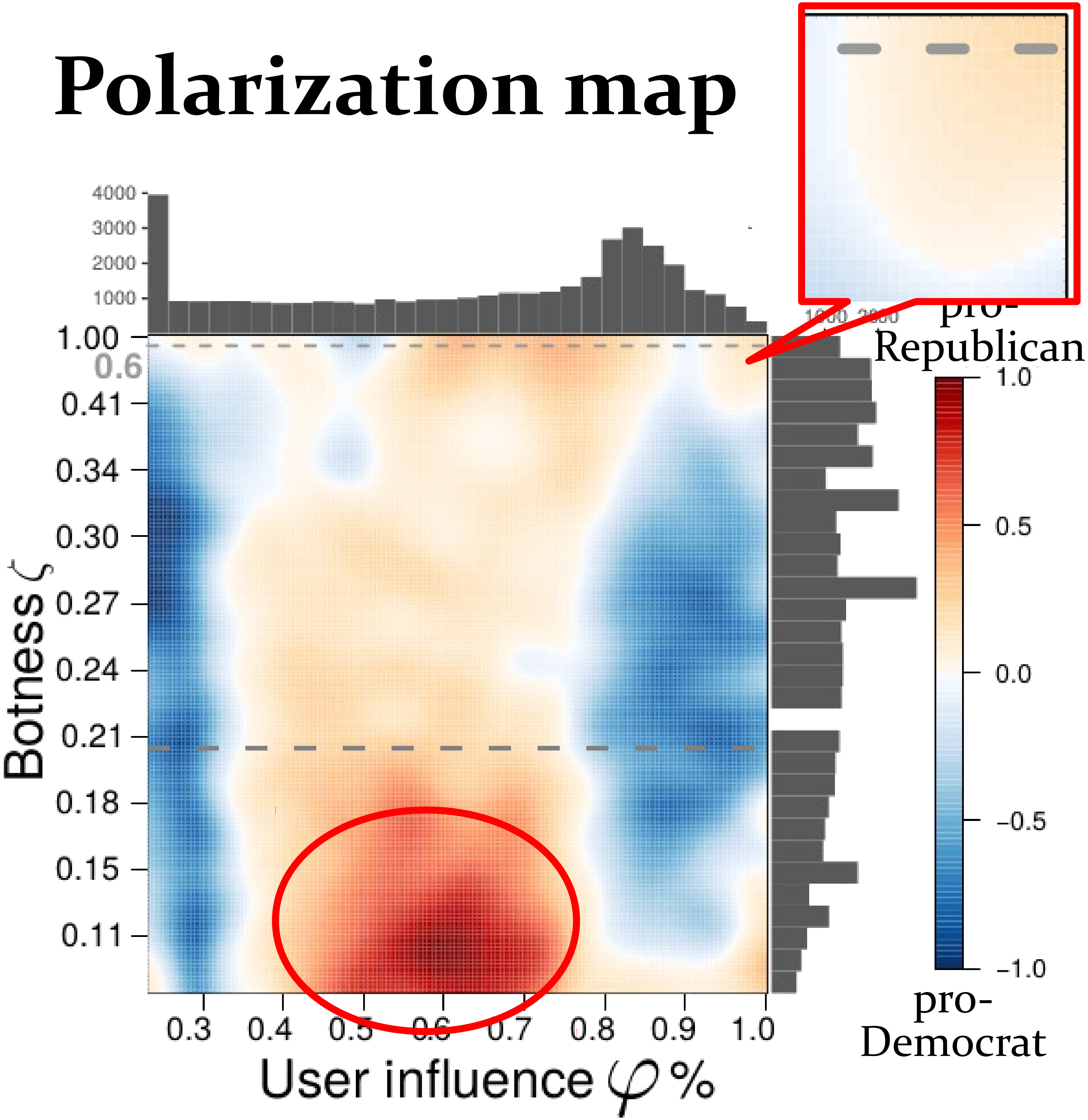
Polarization map



Very highly influential users are pro-Democrat
(**D: 7201**, **R: 5736**)

Highly influential **Bots** are pro-Republican
(**D: 24**, **R: 45**)

Polarization map

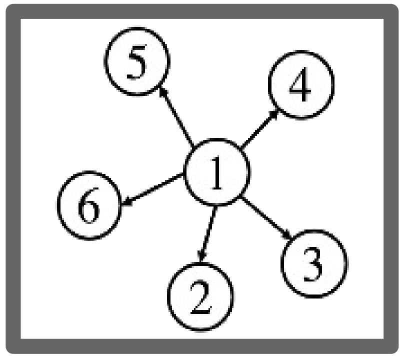


Very highly influential users are pro-Democrat (D: 7201, R: 5736)

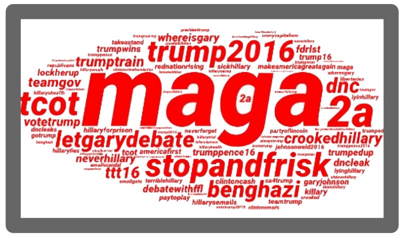
Highly influential **Bots** are pro-Republican (D: 24, R: 45)

Mid-influential humans are pro-Republican (D: 1530, R: 3311)

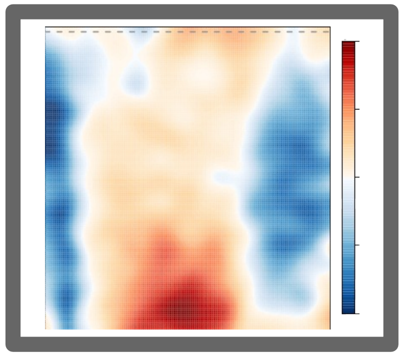
Summary



A scalable algorithm to estimate user influence from latent network structures

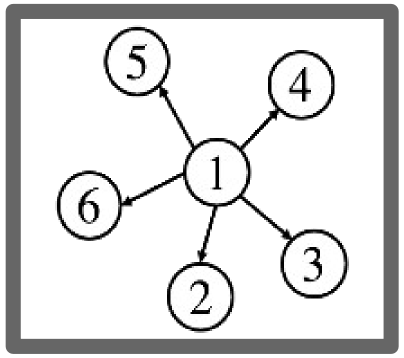


Three measures to quantify the influence, the political partisanship and botness of Twitter users

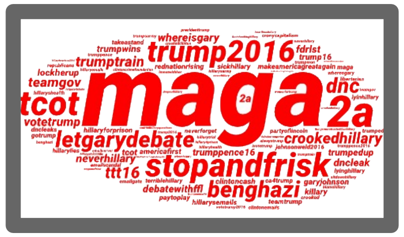


A detailed analysis of the role and influence of socialbots during the first U.S. Presidential debate.

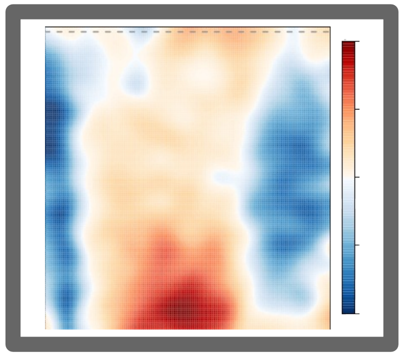
Summary



A scalable algorithm to estimate user influence from latent network structures



Three measures to quantify the influence, the political partisanship and botness of Twitter users



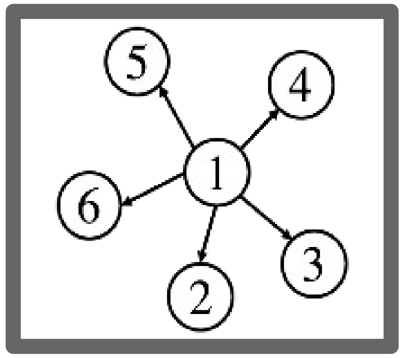
A detailed analysis of the role and influence of socialbots during the first U.S. Presidential debate.

Limitations: Organizational accounts appear as **Bots**; binary partisanship characterization (e.g. independent voters)

Were Bots instrumental for the results of the elections?

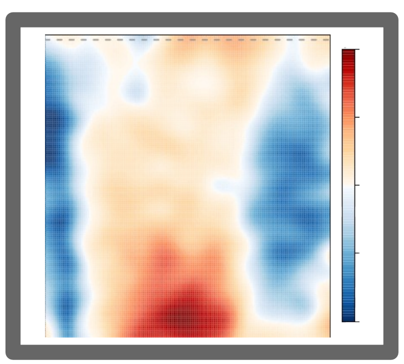
#DebateNight: The Role and Influence of **Socialbots** in the Democratic Process

https://github.com/computationalmedia/cascade-influence



A scalable algorithm to estimate user influence from latent network structures

Three measures to quantify the influence, the political partisanship and botness of Twitter users



A detailed analysis of the role and influence of socialbots during the first U.S. Presidential debate.

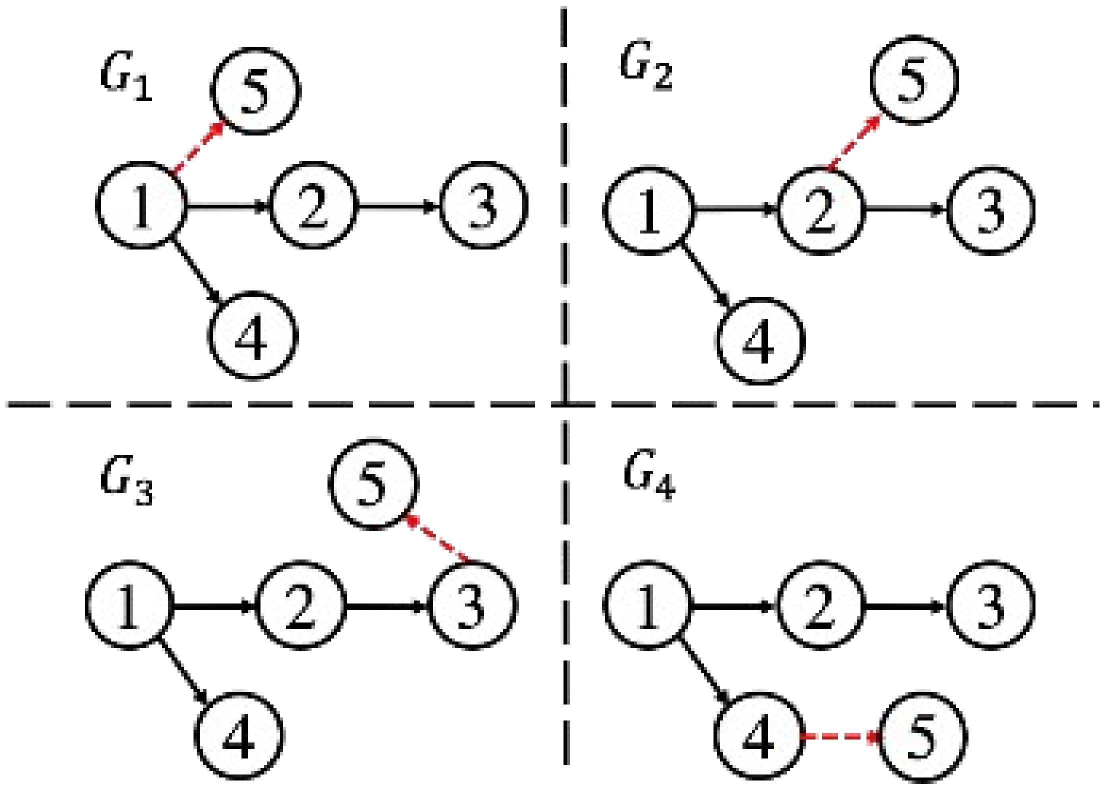
Limitations: Organizational accounts appear as **Bots**; binary partisanship characterization (e.g. independent voters)

Were Bots instrumental for the results of the elections?

Tractable influence computation

Pair-wise influence score:

$$m_{ij} = \begin{cases} \sum_{k=i}^{j-1} m_{ik} p_{kj} & , i < j \\ 1 & , i = j \\ 0 & , i > j. \end{cases}$$



Influence of a tweet:

$$\varphi(v_i) = \sum_{j=1}^{k} m_{ij}.$$

Influence of a user:

$$\varphi(u) = \frac{\sum_{v \in \mathcal{T}(u)} \varphi(v)}{|\mathcal{T}(u)|}, \mathcal{T}(u) = \{v | u_v = u\}$$