#DebateNight - Role of Twitter Socialbots During US Presidential Debate

Marian-Andrei Rizoiu

Behavioral Data Science

UTS
UNIVERSITY OF TECHNOLOGY SYDNEY

The research group

1 research associate, 3 PhD students, 2 Honors students, 1 lecturer







Research income & grants

Behavioral Data Science

~$460k

2019 – current:	Crawford School of Public Policy grants, "**Evaluating democratic equity through analysing data around public donation to presidential candidates**", co-CI.
2019 – current:	UTS cross-faculty collaboration scheme, "**SocialSense: Making sense of the opinions and interactions of online users**", CI.
2019 – current:	Data61 Challenge model grants, "**Adaptive skills taxonomy to enable labour market agility"**, CI.
2018	ANU Social Science Cross-College Grants, **"Advanced tools and methods for analysing the role and influence of bots in social media"**, CI.
2018	ANU Social Science Cross-College Grants, **"Identify Hate Speech and Predict Mass Atrocities"**, CI.

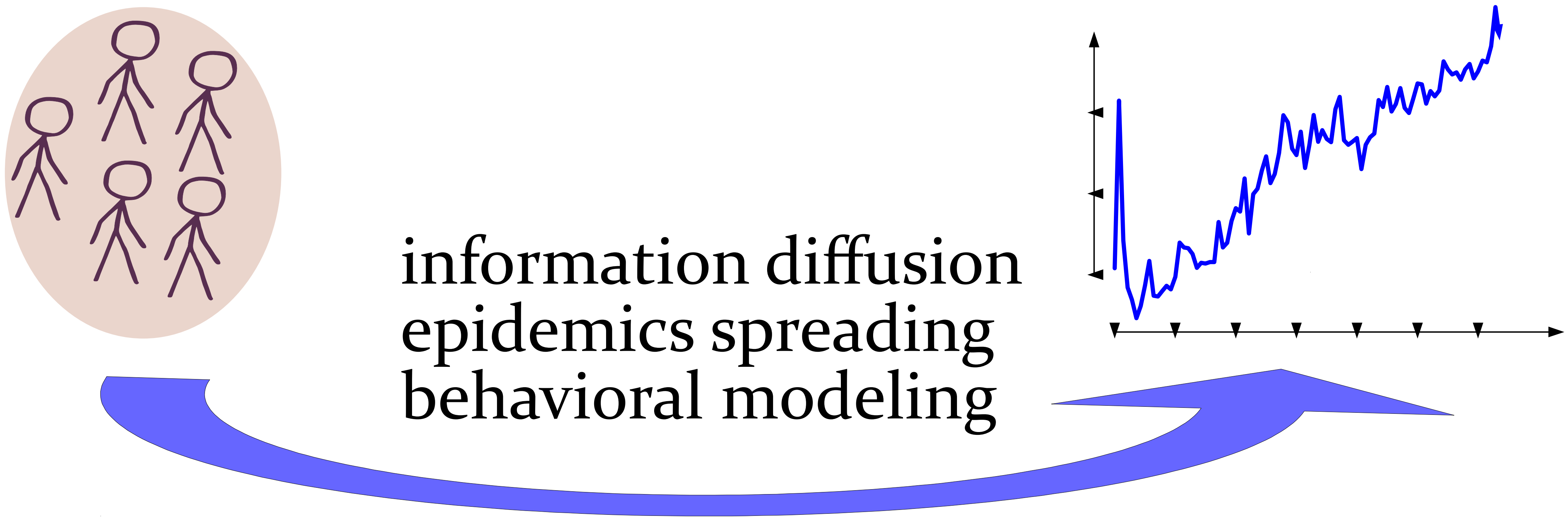
WIKIMEDIA
FOUNDATION

NETFLIX

Research objectives

Behavioral
Data Science

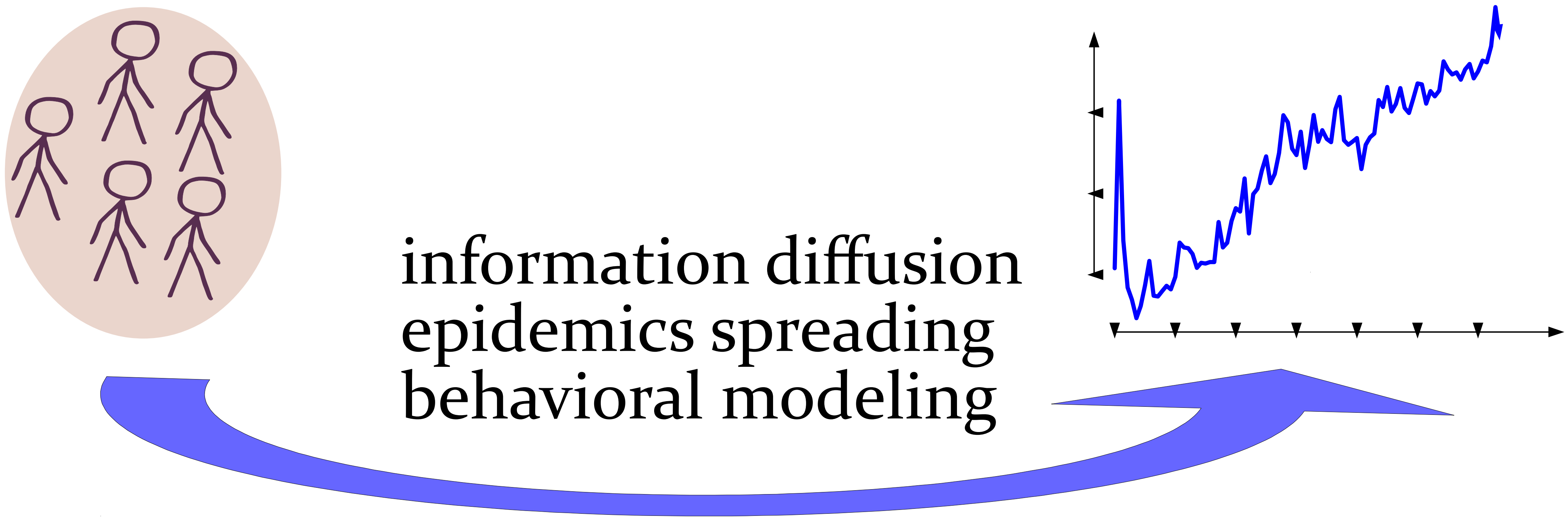
1.



Research objectives

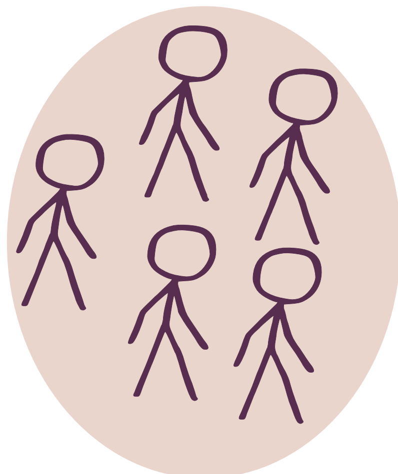
Behavioral
Data Science

1.

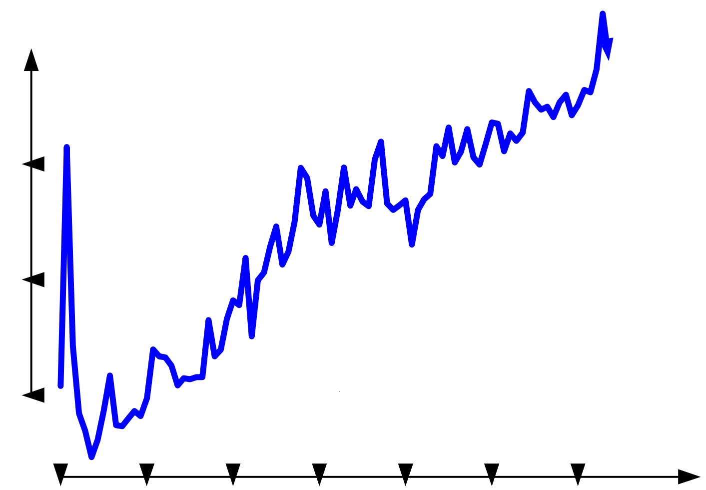


Research objectives

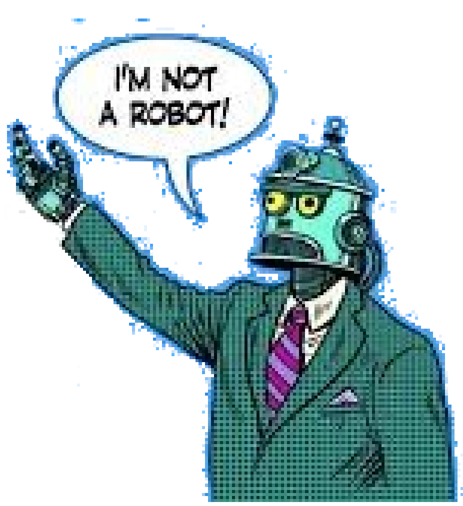
1.



information diffusion
epidemics spreading
behavioral modeling



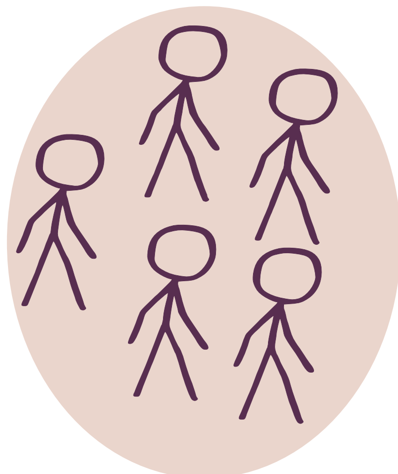
2.



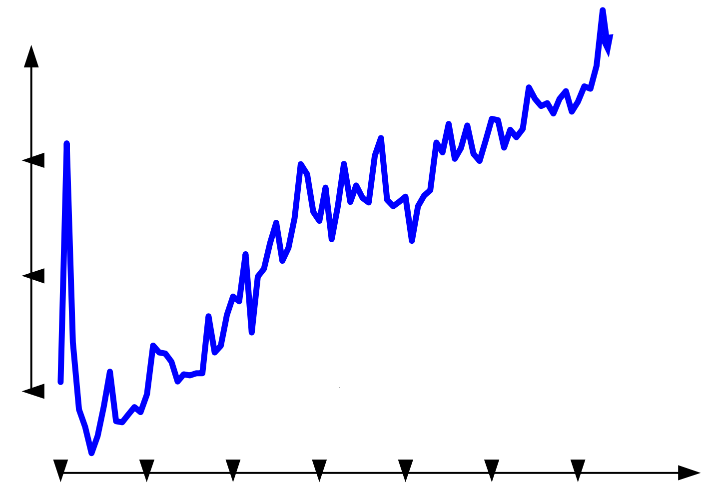
Research objectives

Behavioral Data Science

1.

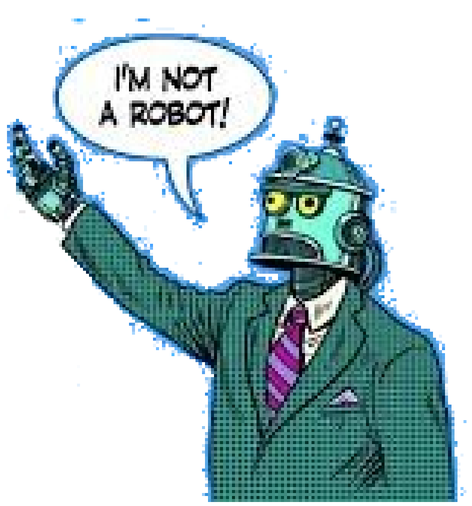


information diffusion
epidemics spreading
behavioral modeling



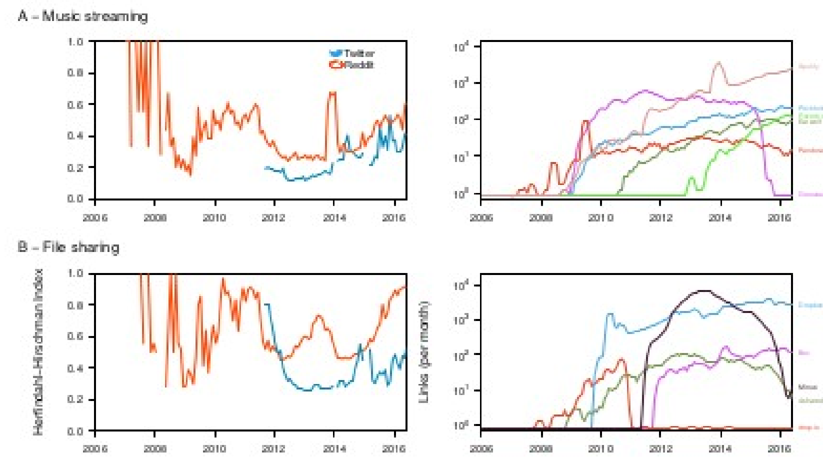
3.

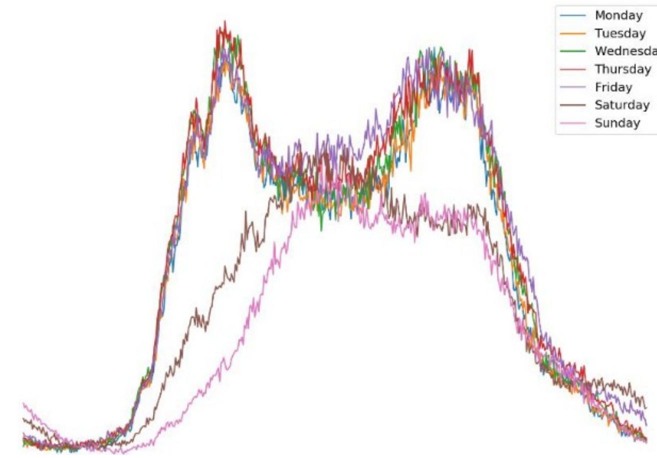
2.



Other projects

Behavioral Data Science

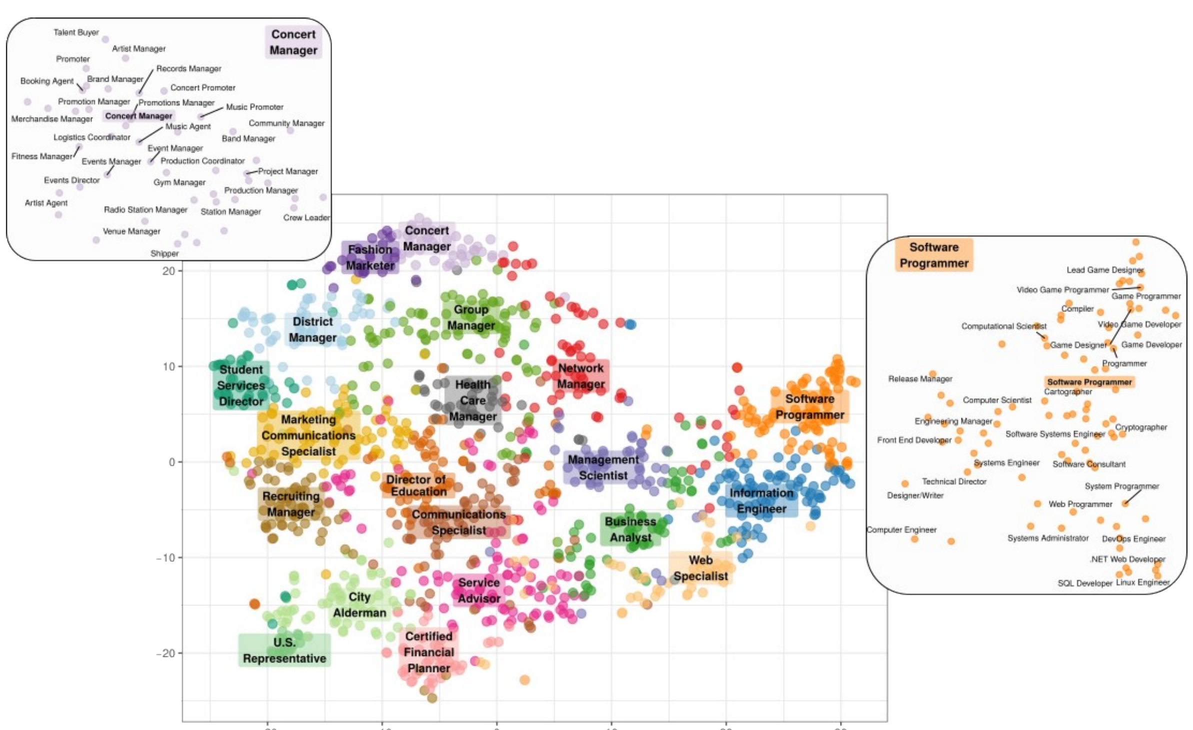


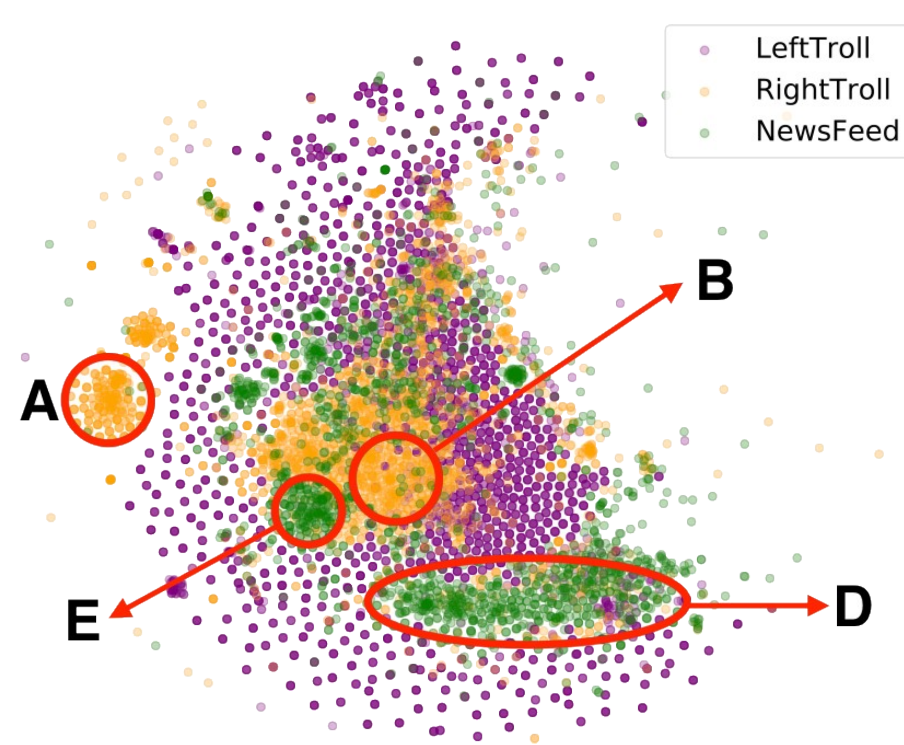


Wikipedia privacy

Online Diversity

Smart traffic

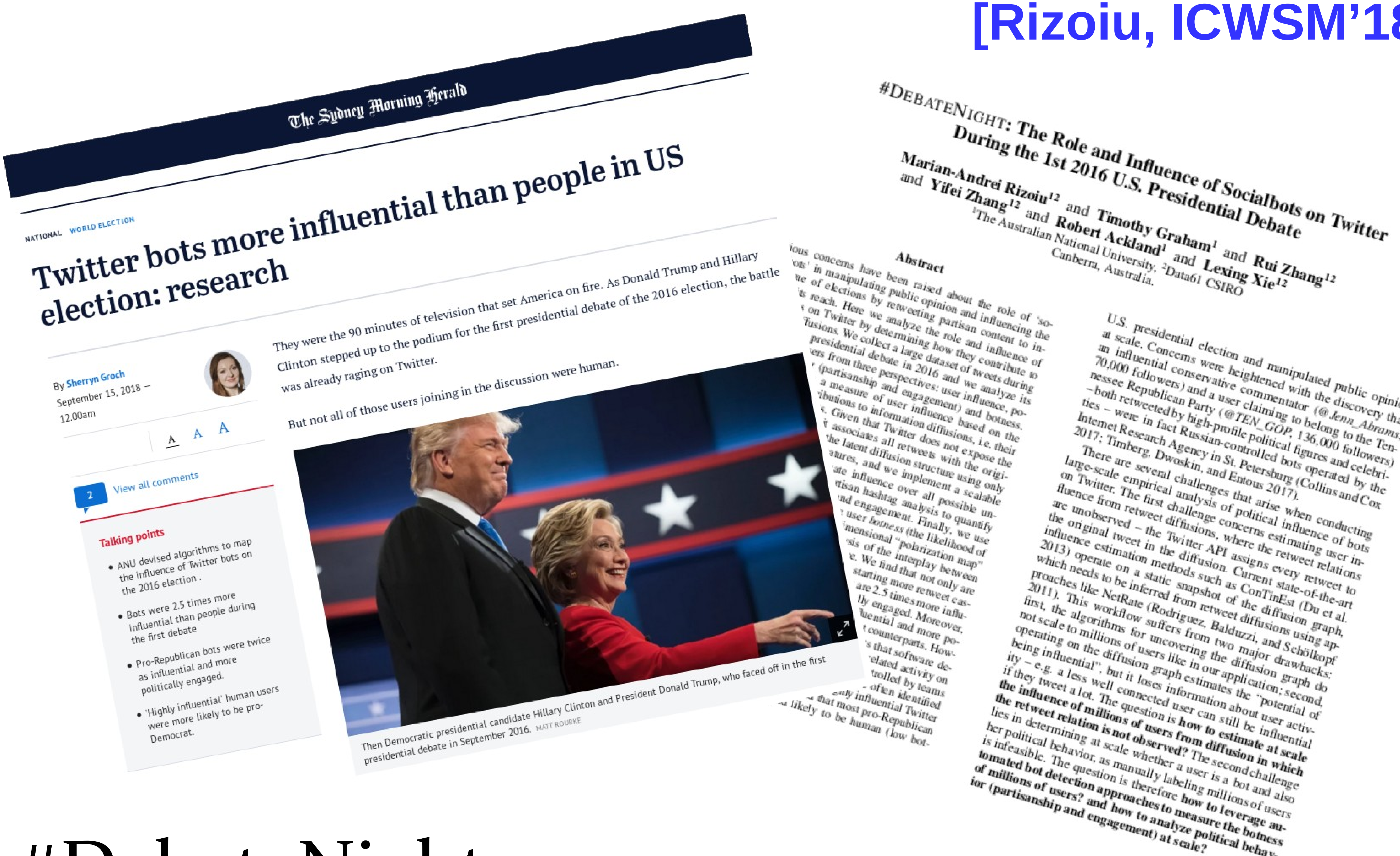




Vocation compass

Busting Russian Trolls

[Rizoiu, ICWSM'18]



The Sydney Morning Herald

NATIONAL WORLD ELECTION

Twitter bots more influential than people in US election: research

By Sherryn Groch
September 15, 2018 — 12.00am

They were the 90 minutes of television that set America on fire. As Donald Trump and Hillary Clinton stepped up to the podium for the first presidential debate of the 2016 election, the battle was already raging on Twitter.

But not all of those users joining in the discussion were human.

Then Democratic presidential candidate Hillary Clinton and President Donald Trump, who faced off in the first presidential debate in September 2016. MATT ROURKE

2 View all comments

Talking points

- ANU devised algorithms to map the influence of Twitter bots on the 2016 election .
- Bots were 2.5 times more influential than people during the first debate
- Pro-Republican bots were twice as influential and more politically engaged.
- 'Highly influential' human users were more likely to be pro-Democrat.

#DebateNight: The Role and Influence of Socialbots on Twitter During the 1st 2016 U.S. Presidential Debate

Marian-Andrei Rizoiu[1,2] and Timothy Graham[1] and Rui Zhang[1,2] and Yifei Zhang[1,2] and Robert Ackland[1] and Lexing Xie[1,2]
[1]The Australian National University, [2]Data61 CSIRO
Canberra, Australia

U.S. presidential election and manipulated public opinion at scale. Concerns were heightened with the discovery that an influential conservative commentator (@Jenn_Abrams, 70,000 followers) and a user claiming to belong to the Tennessee Republican Party (@TEN_GOP, 136,000 followers) – both retweeted by high-profile political figures and celebrities – were in fact Russian-controlled bots operated by the Internet Research Agency in St. Petersburg (Collins and Cox 2017; Timberg, Dwoskin, and Entous 2017).

There are several challenges that arise when conducting large-scale empirical analysis of political influence of bots on Twitter. The first challenge concerns estimating user influence from retweet diffusions, where the retweet relations are unobserved – the Twitter API assigns every retweet to the original tweet in the diffusion. Current state-of-the-art influence estimation methods such as ConTinEst (Du et al. 2013) operate on a static snapshot of the diffusion graph, which needs to be inferred from retweet diffusions using approaches like NetRate (Rodriguez, Balduzzi, and Schölkopf 2011). This workflow suffers from two major drawbacks: first, the algorithms for uncovering the diffusion graph do not scale to millions of users like in our application; second, operating on the diffusion graph estimates the "potential of being influential", but it loses information about user activity – e.g. a less well connected user can still be influential if they tweet a lot. The question is **how to estimate at scale the influence of millions of users from diffusion in which the retweet relation is not observed?** The second challenge lies in determining at scale whether a user is a bot and also her political behavior, as manually labeling millions of users is infeasible. The question is therefore **how to leverage automated bot detection approaches to measure the botness of millions of users? and how to analyze political behavior (partisanship and engagement) at scale?**

#DebateNight
Role of Twitter Socialbots
During US Presidential Debate

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

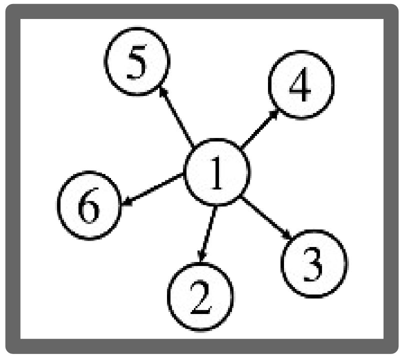
Dataset stats:

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

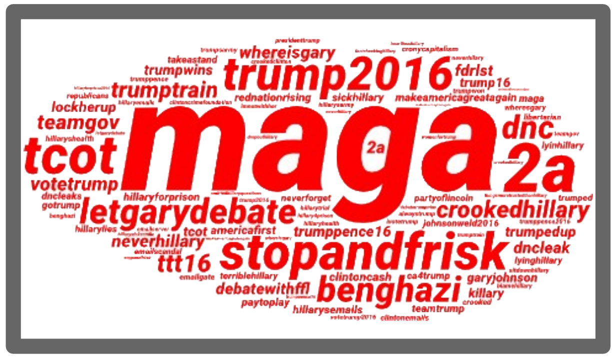
Hashtags:

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

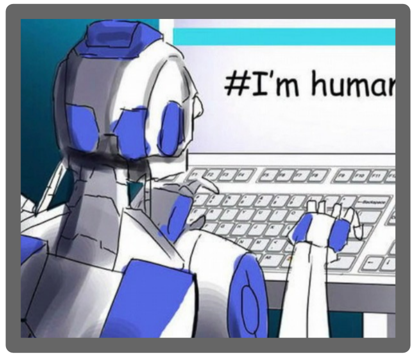
Presentation outline



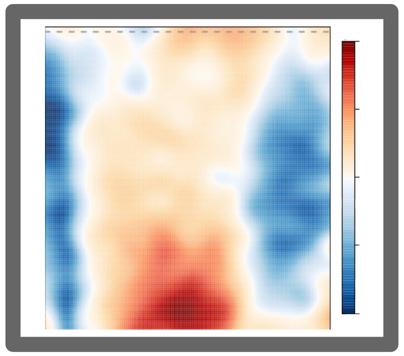
User influence



Political partisanship

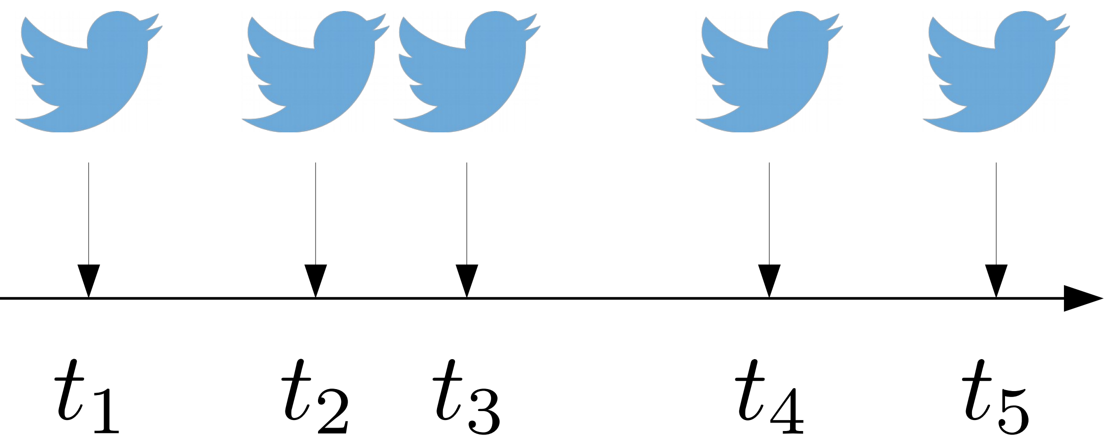


User botness

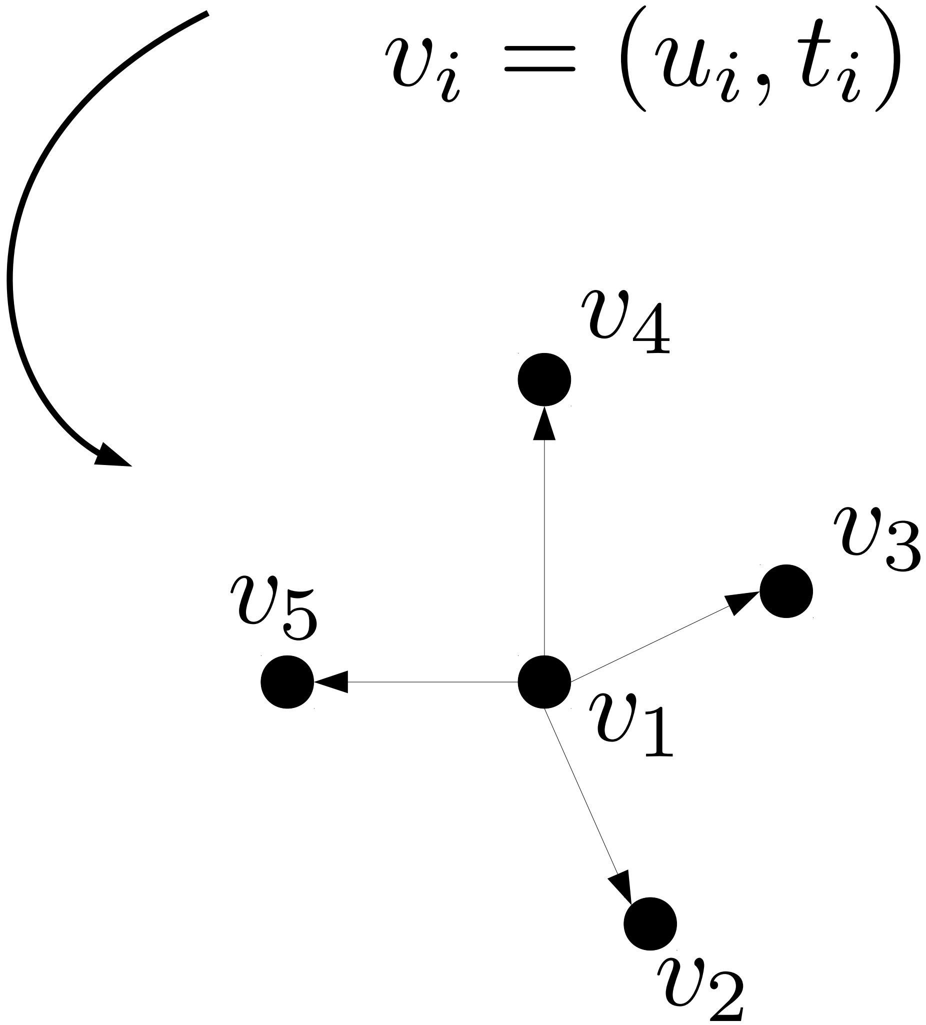


Analyze political
behavior of bots

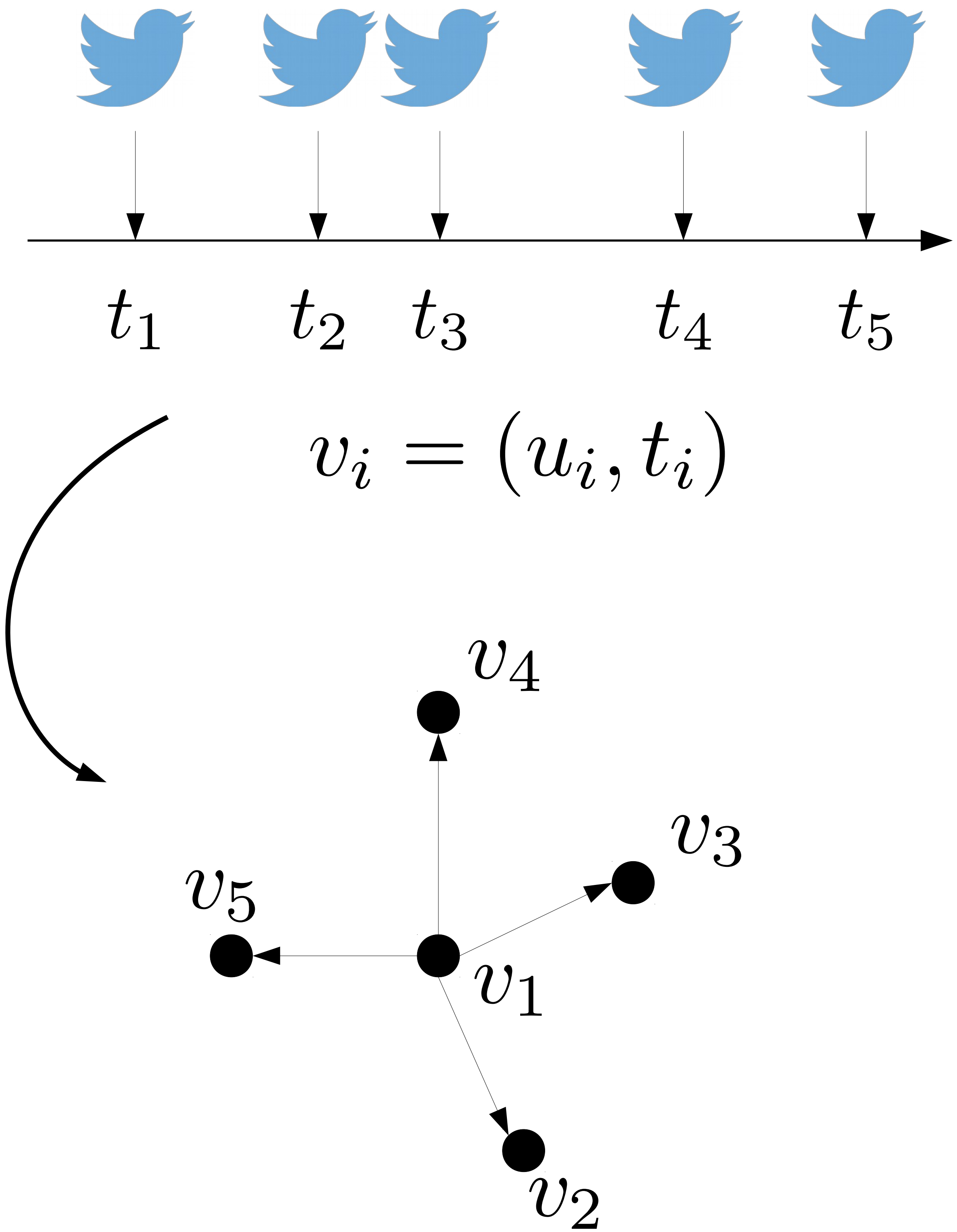
Retweet influence (1)



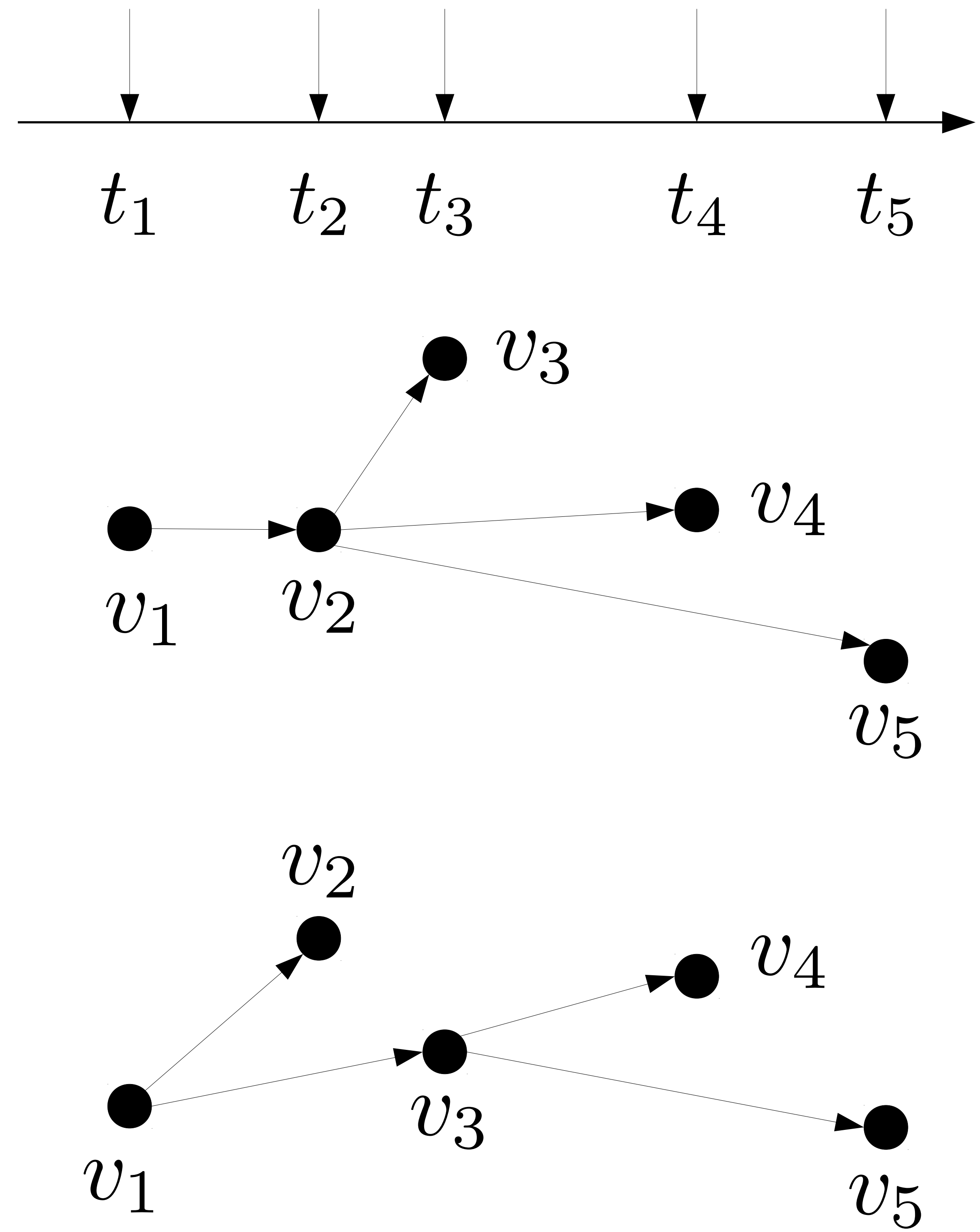
$$v_i = (u_i, t_i)$$



Retweet influence (1)

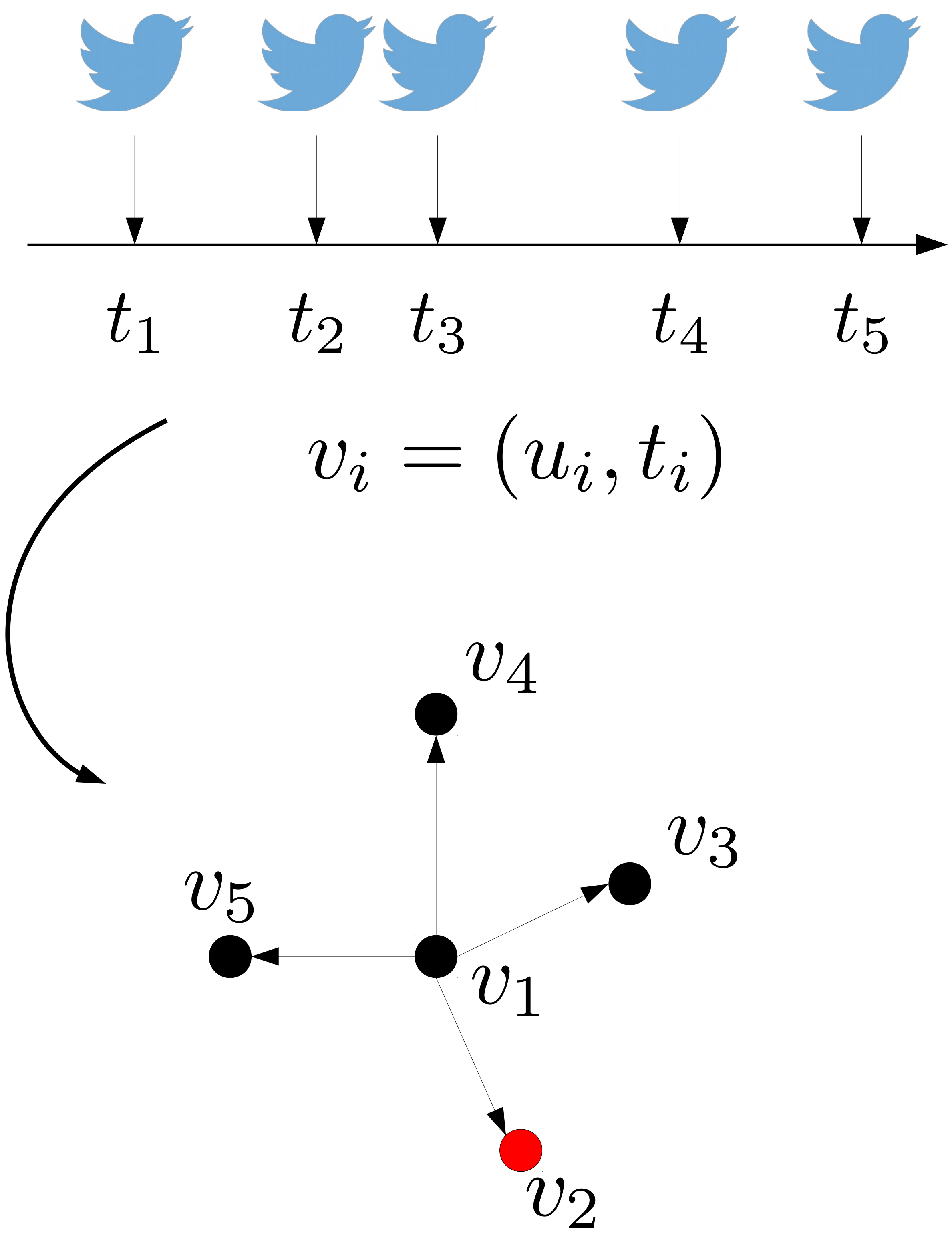


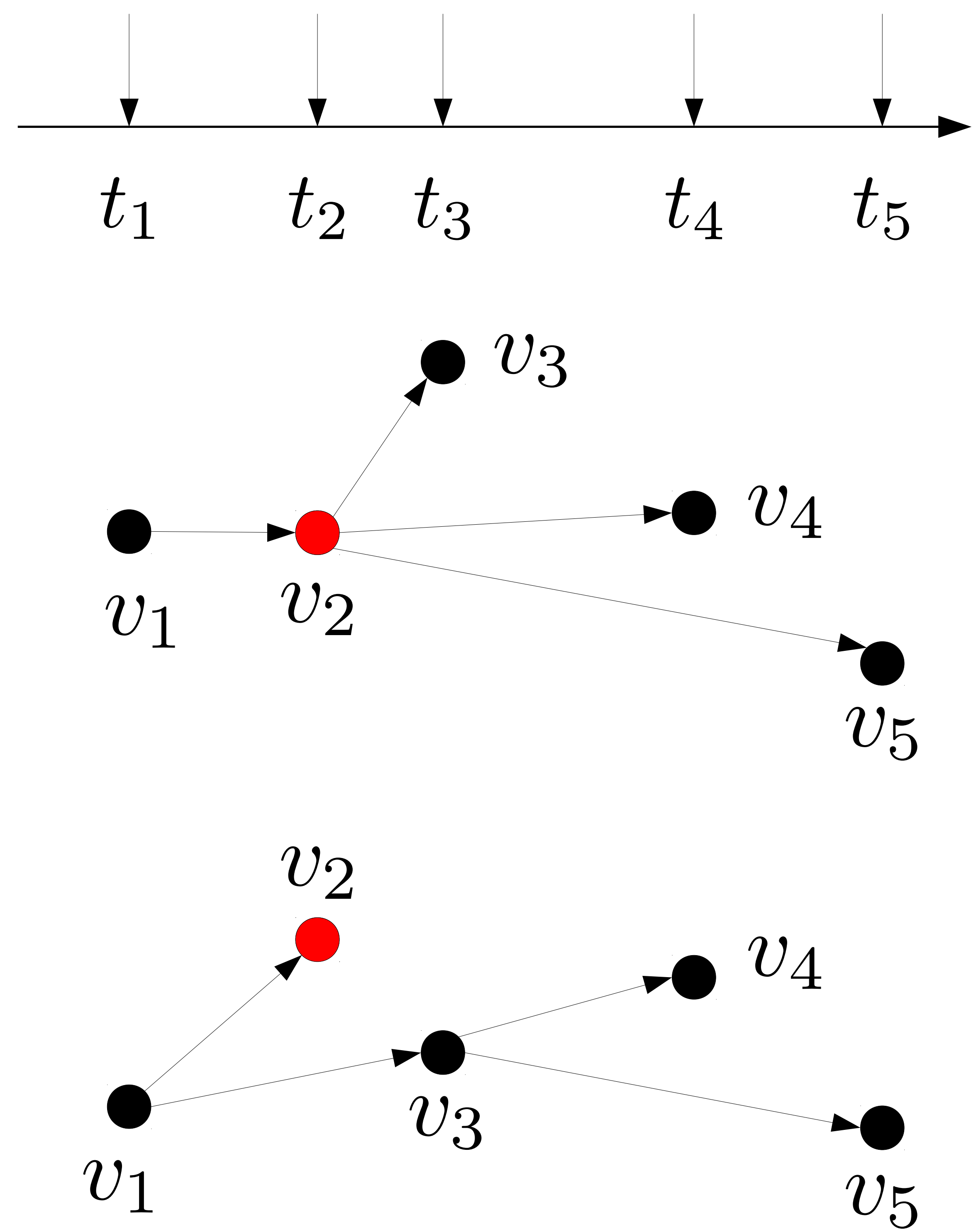
Diffusion trees and influence



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

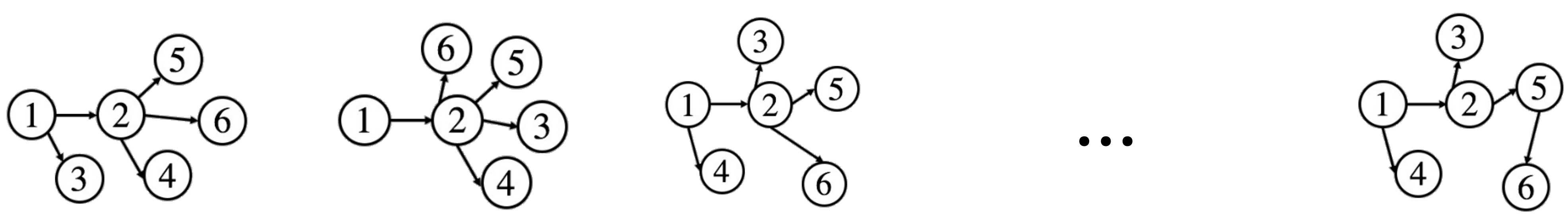
$$p_{ij} = \frac{\mathbf{m_i e^{-r(t_j - t_i)}}}{\sum_{k=1}^{j-1} m_k e^{-r(t_j - t_k)}}$$

branching probability

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

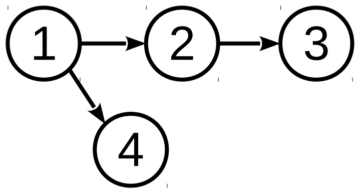
Tweet influence: the expected number of retweets, averaged over all possible trees.

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

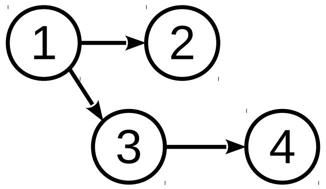


Tractable influence computation

Pair-wise influence score m_{ij}



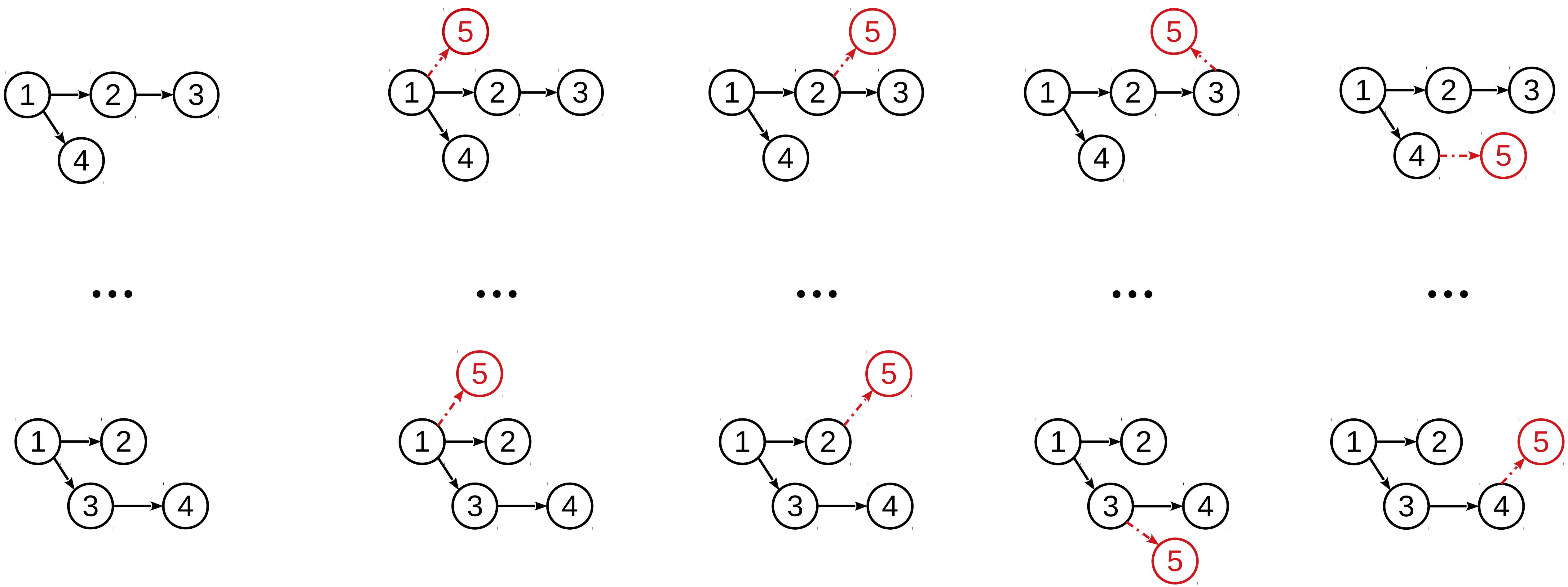
...



Tractable influence computation

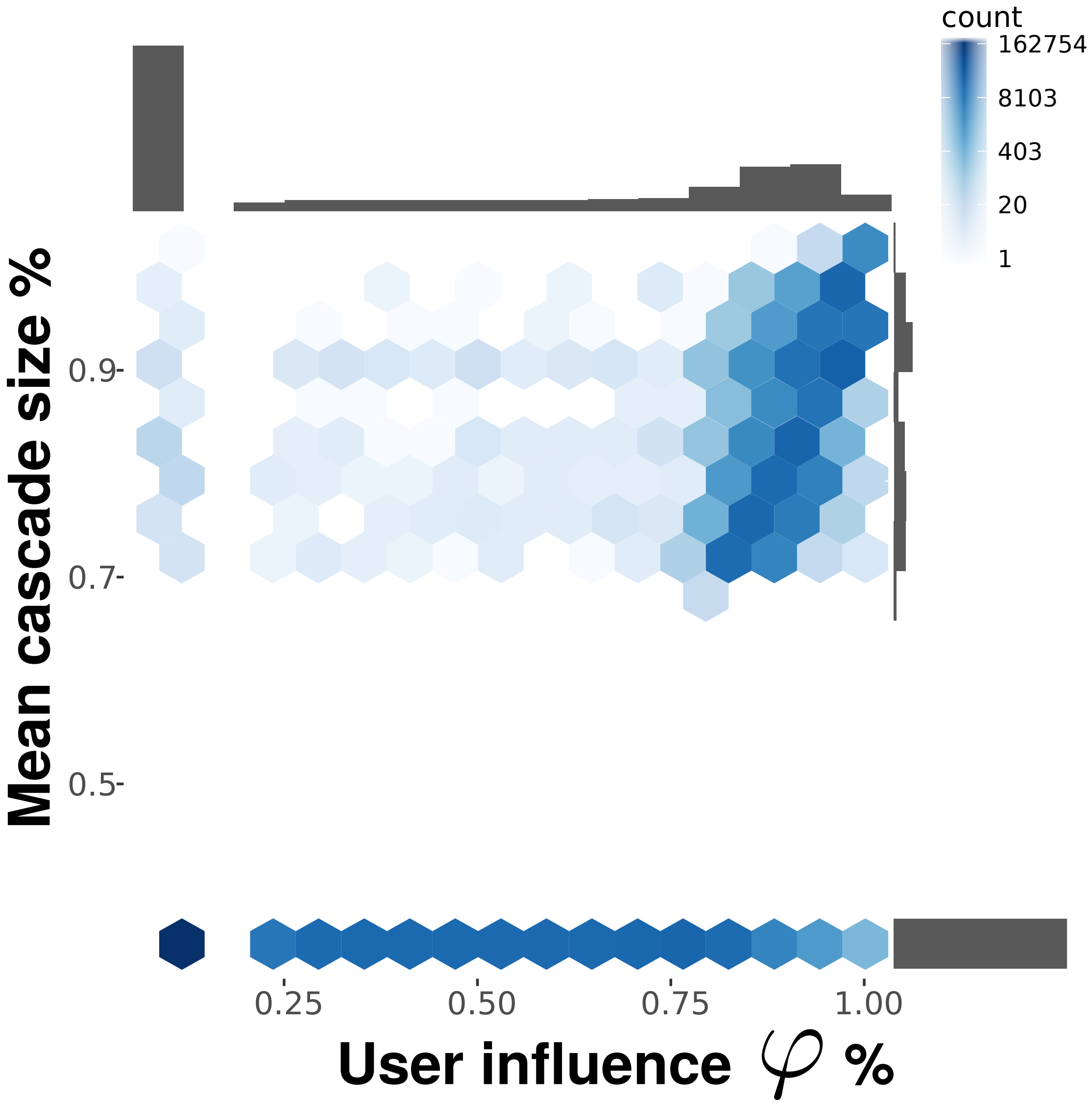
Pair-wise influence score m_{ij}

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



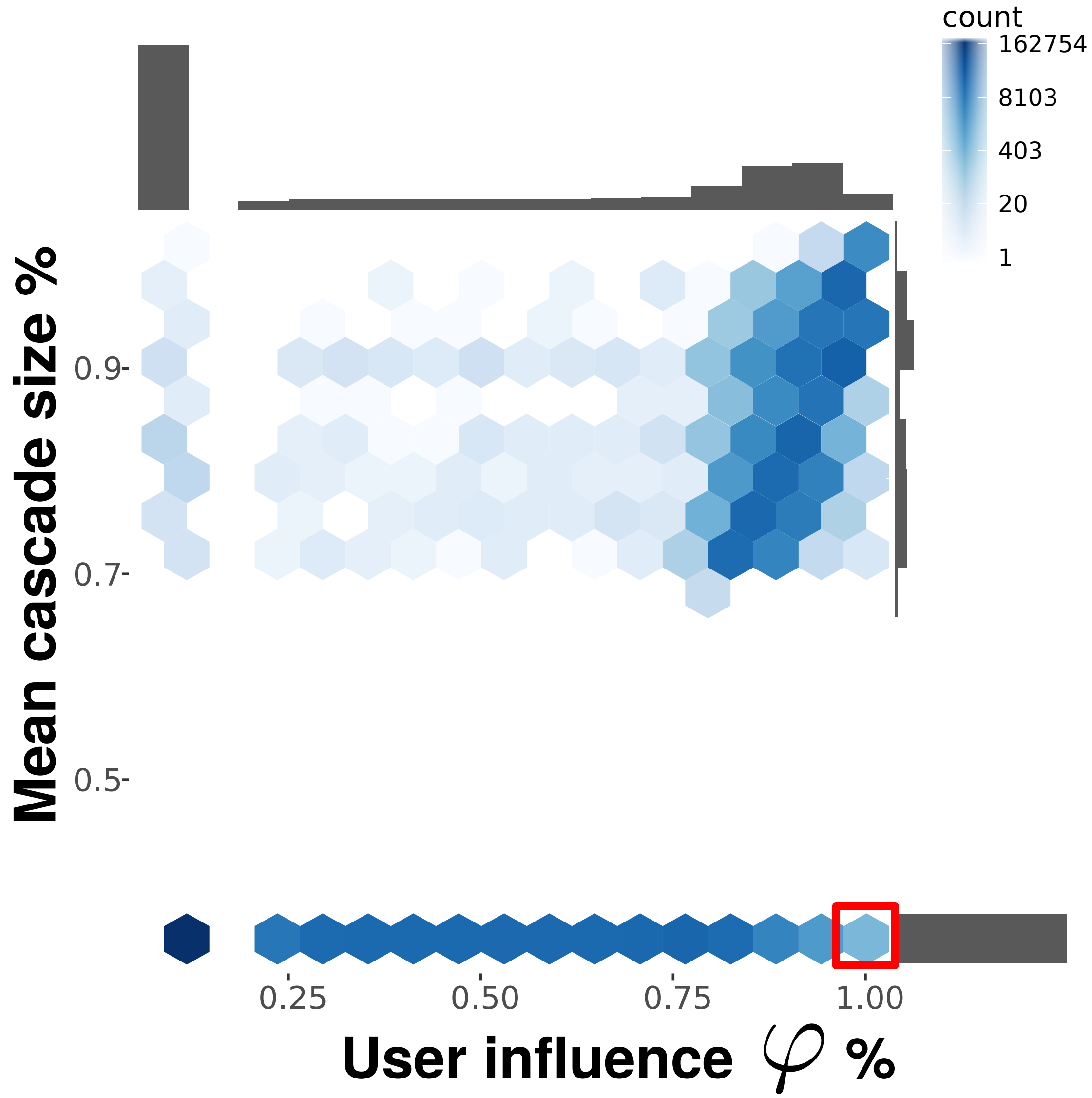
Recursive algorithm $\quad O(n^3)$

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)

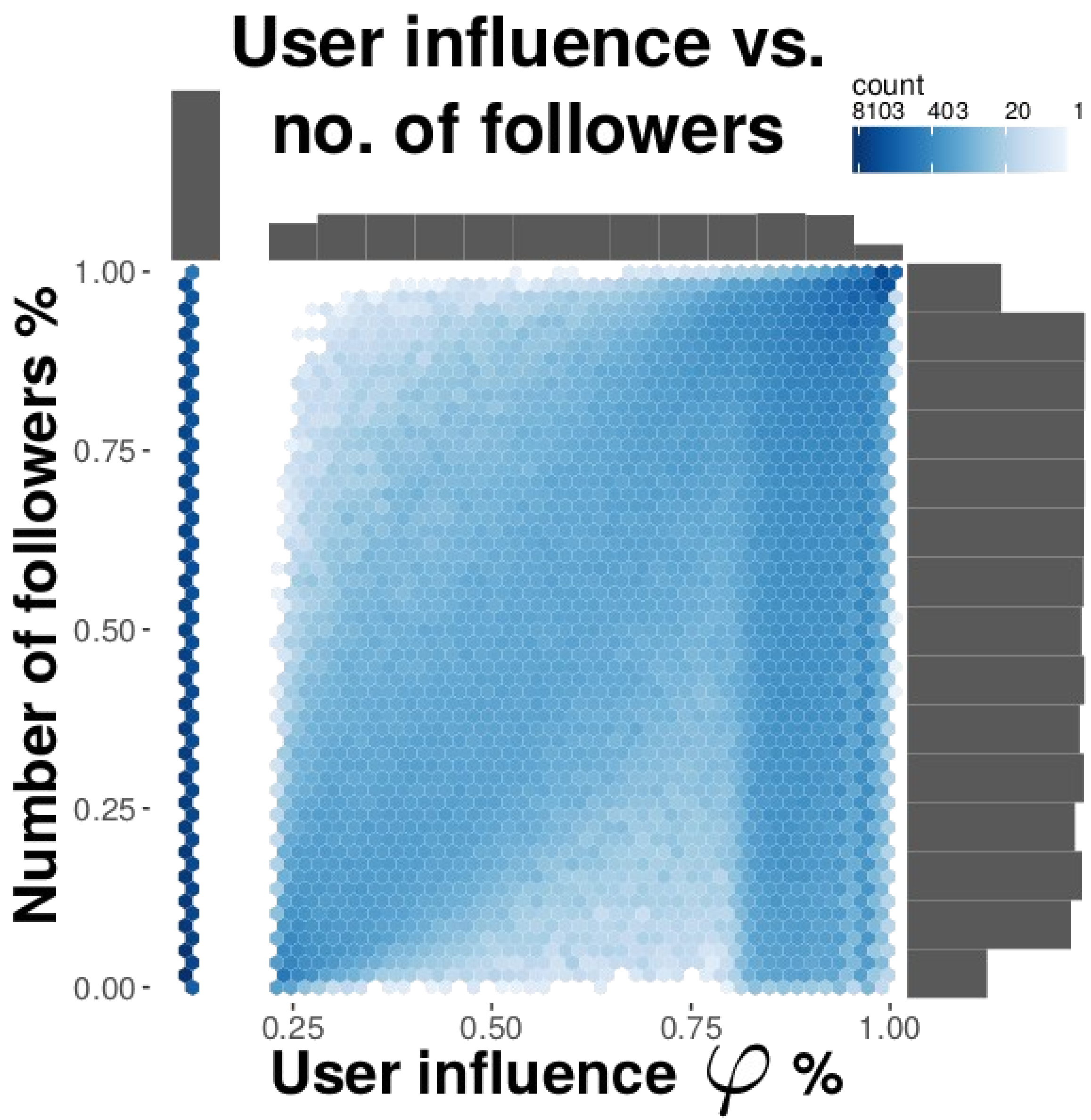
actor and filmmaker

10.8 million followers

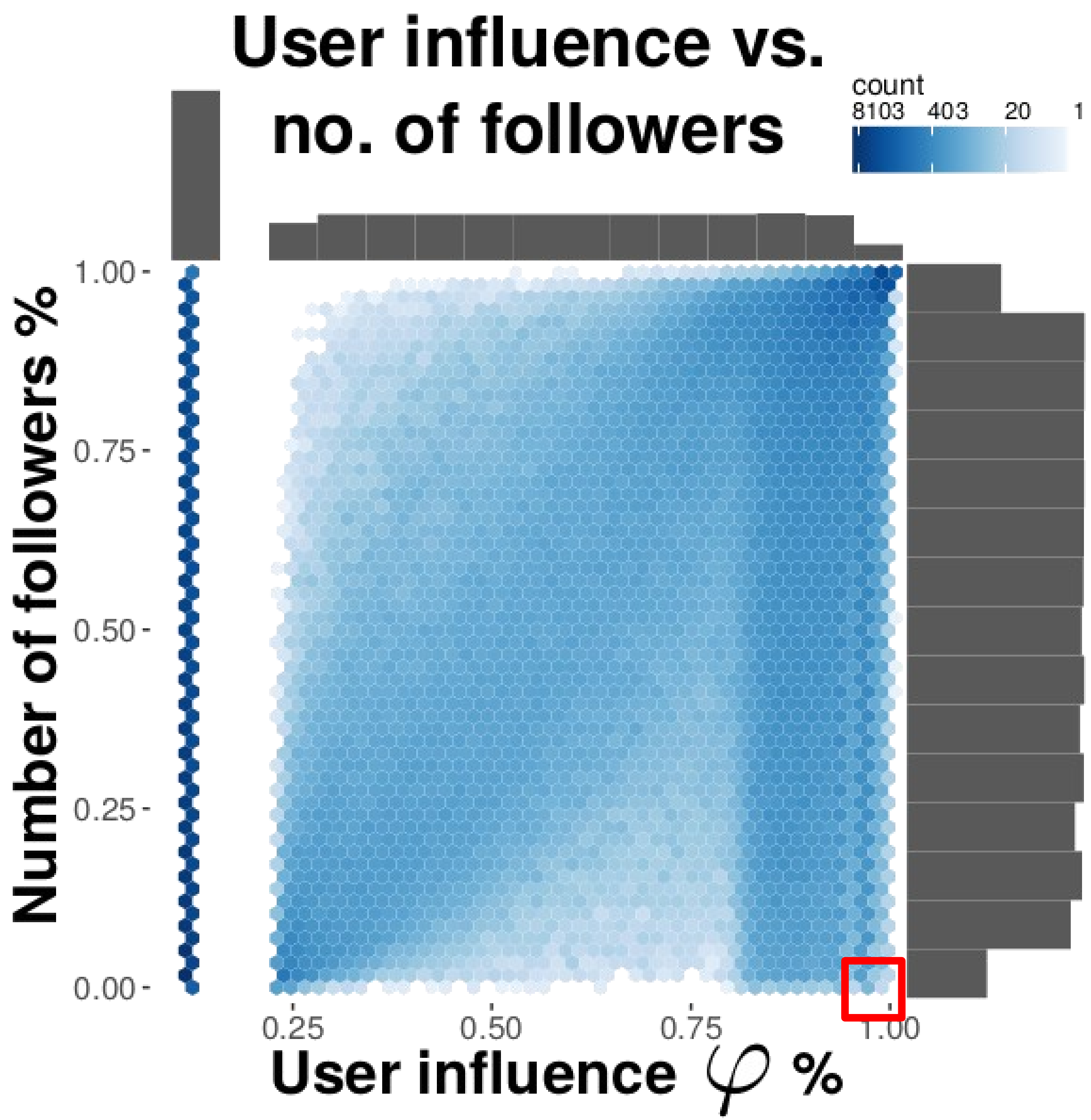
comedian

2.1 million followers

Influence vs. number of followers



Influence vs. number of followers



now suspended

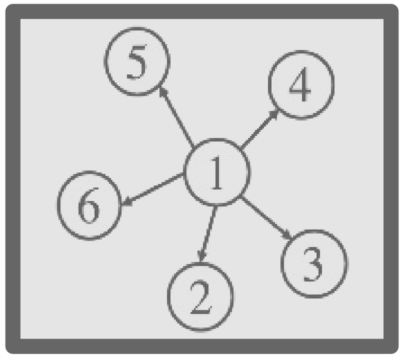
2 followers

1 follower

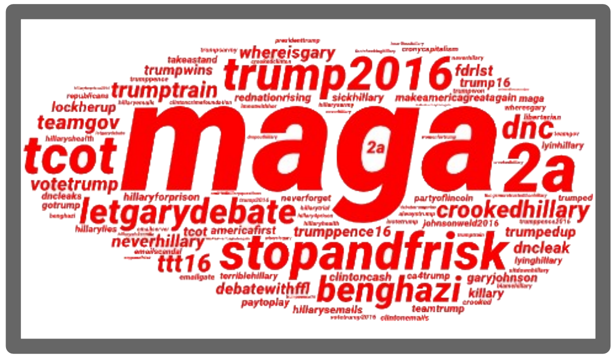
Initiated a big cascade

Initiated a big cascade

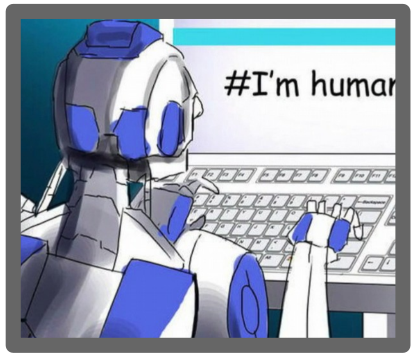
Presentation outline



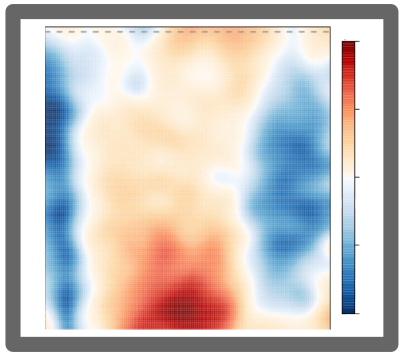
User influence



Political partisanship



User botness

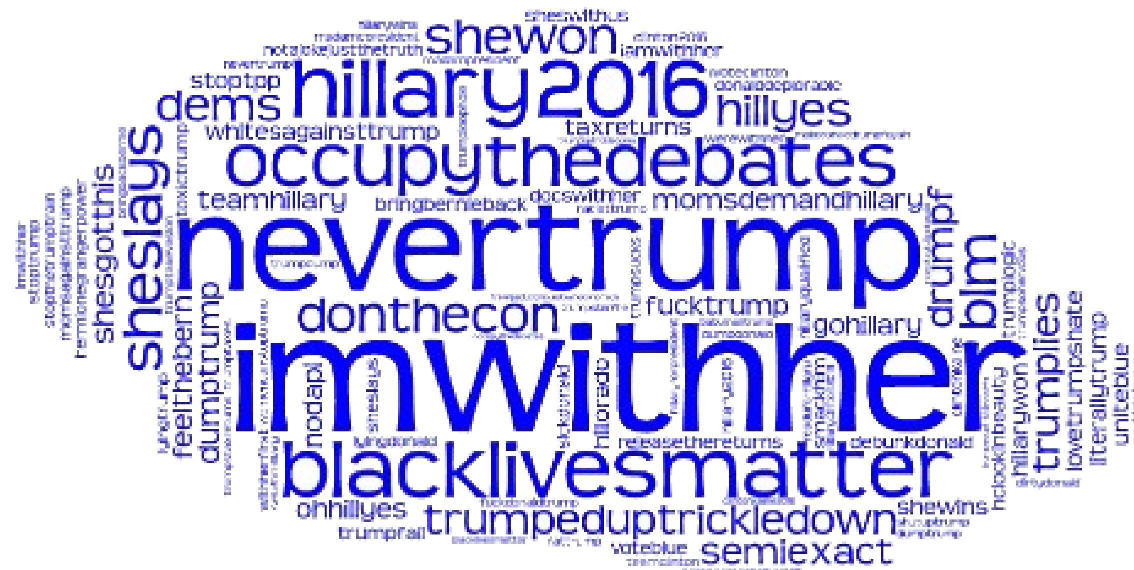


Analyze political behavior of bots

Political polarization (1)

Protocol:

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



Partisanship stats:

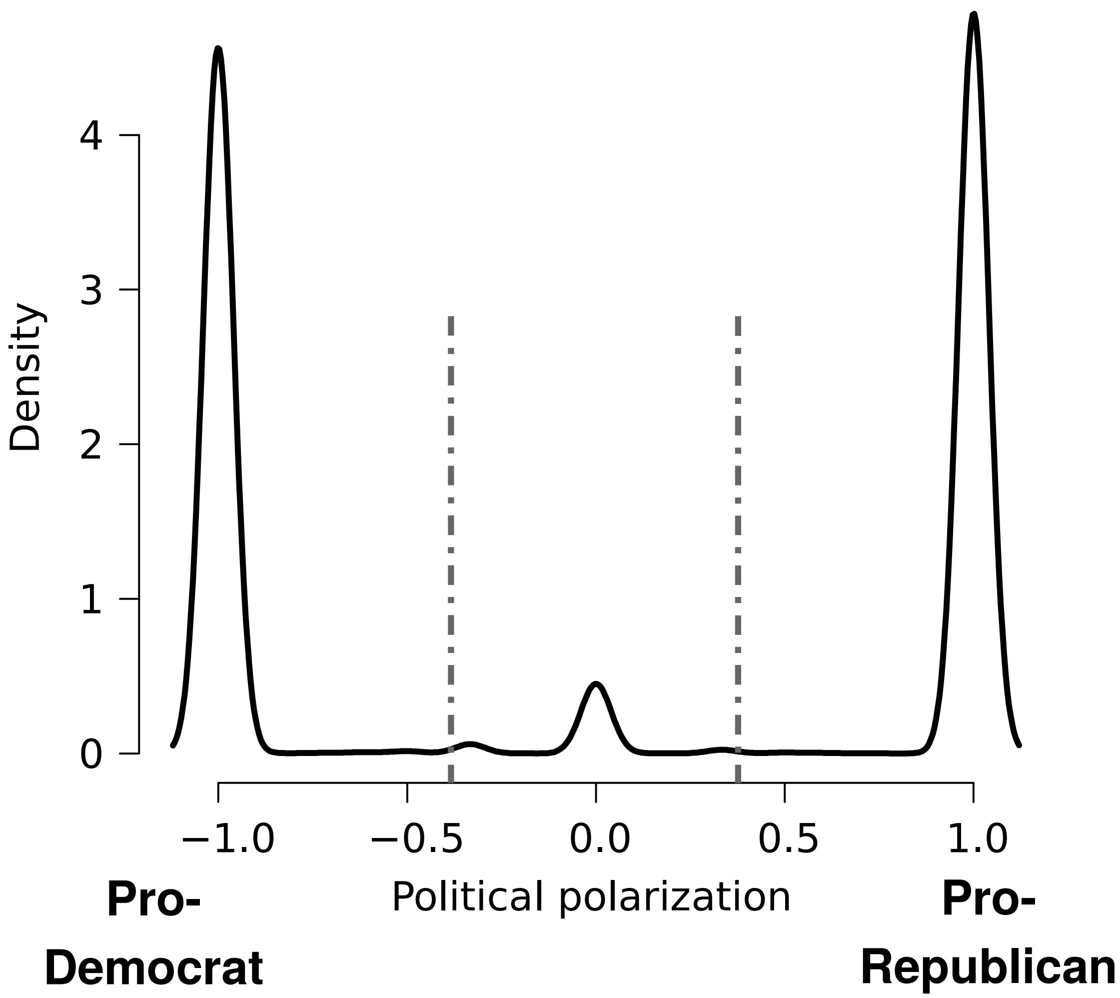
- pro-Democrat hashtags: **93**
- pro-Republican hashtags: **86**
- partisan tweets: **65K**
- partisan users: **47K**

Political polarization (2)

For each user *i*:

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

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

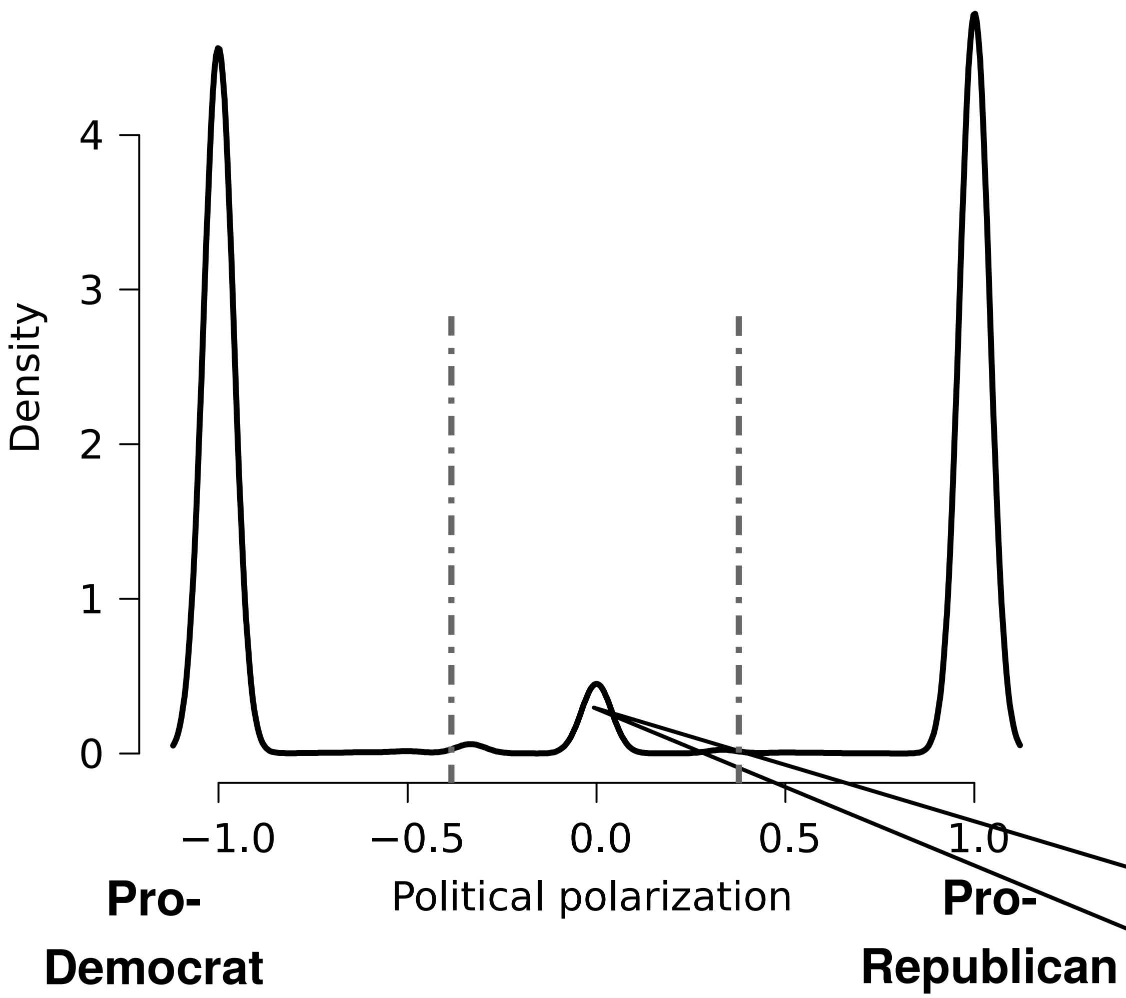


Political polarization (2)

For each user *i*:

- *dem*$_i$ – #democrat hashtags
- *rep*$_i$ – #republican hashtags

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



Let's Get READY TO RUMBLE AND TELL LIES. #debatenight #debates #Debates2016 #cnn #nevertrump #neverhillary #Obama

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

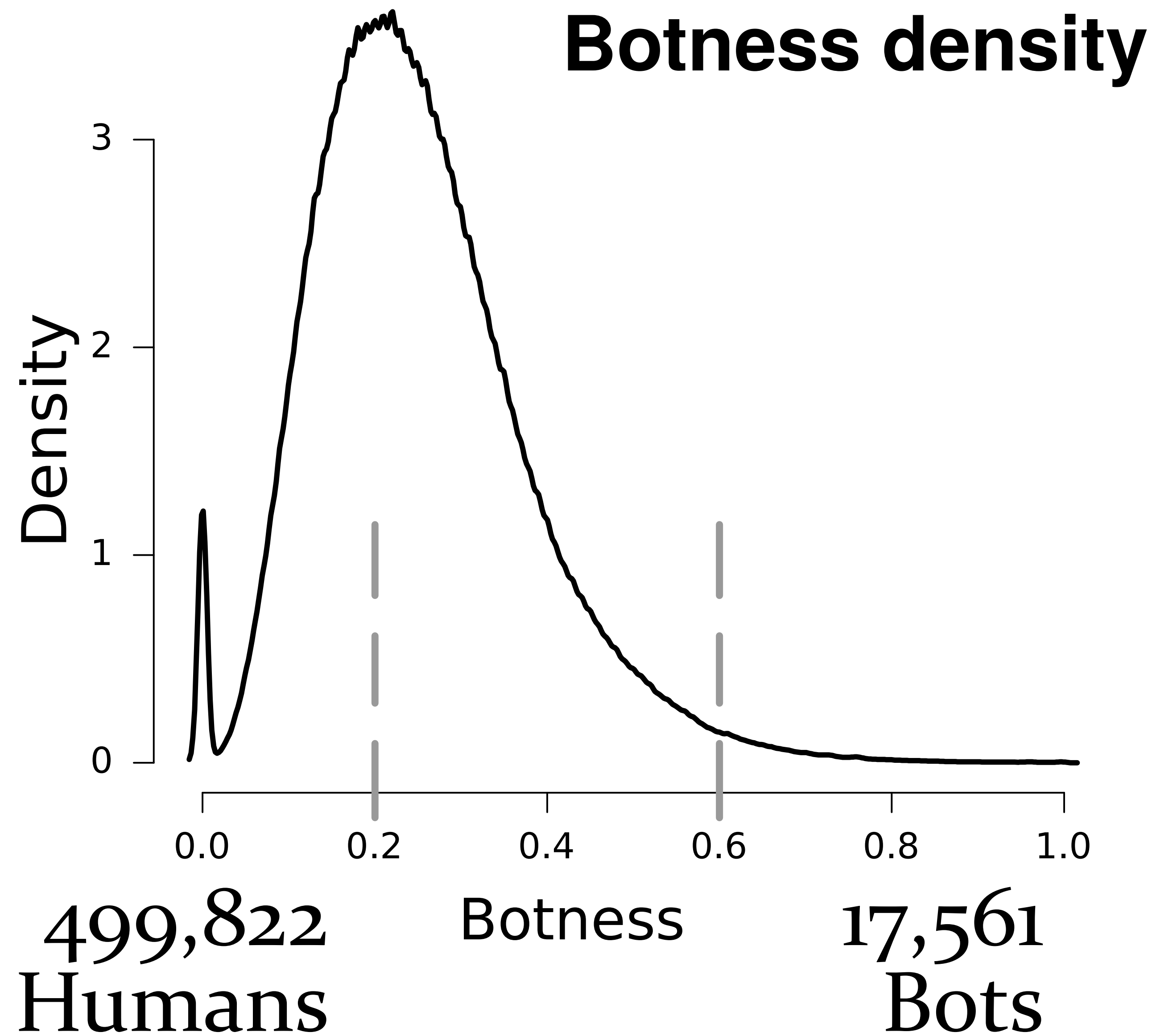
Three populations

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

Separating bots from humans

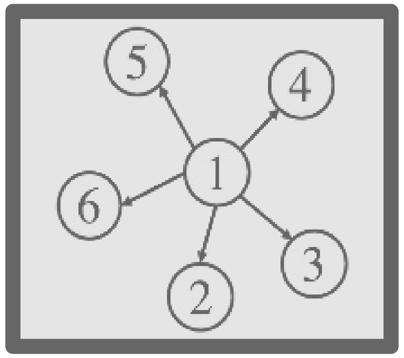
Three populations

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

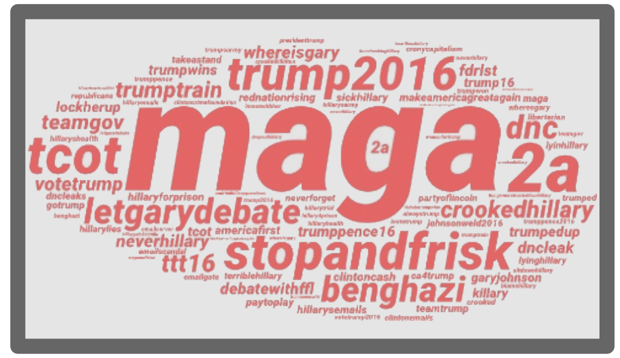


[Varol et al, ICWSM'17] use a threshold of 0.5

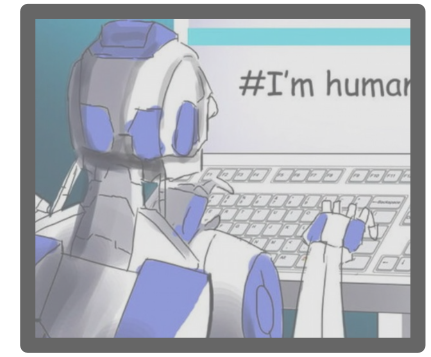
Presentation outline



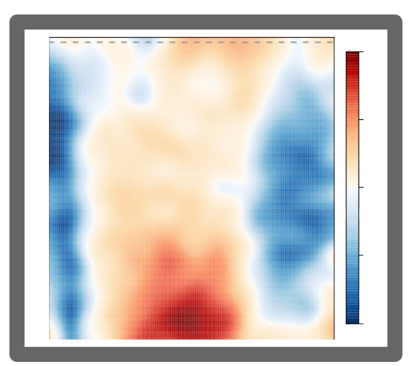
User influence



Political partisanship

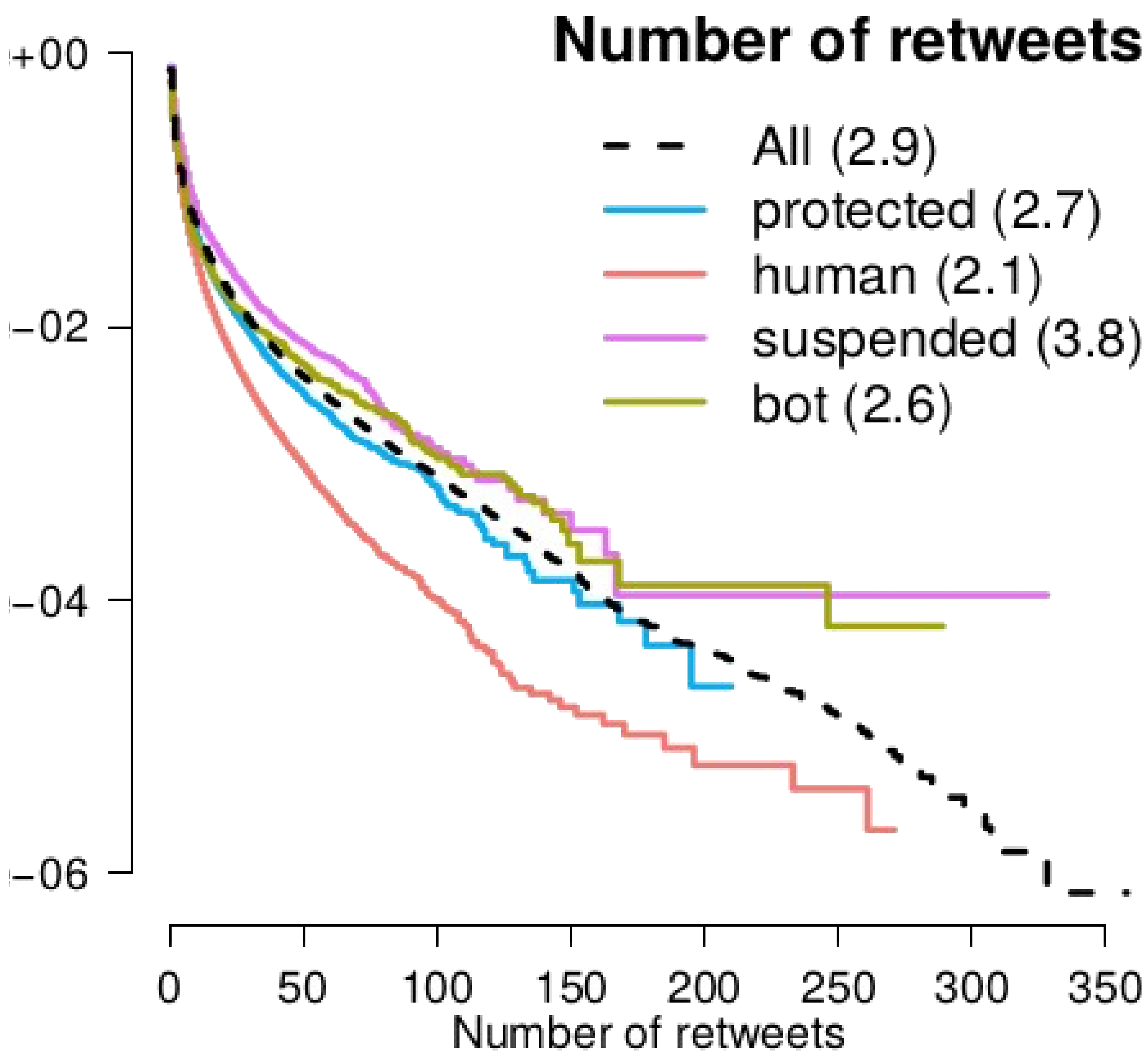


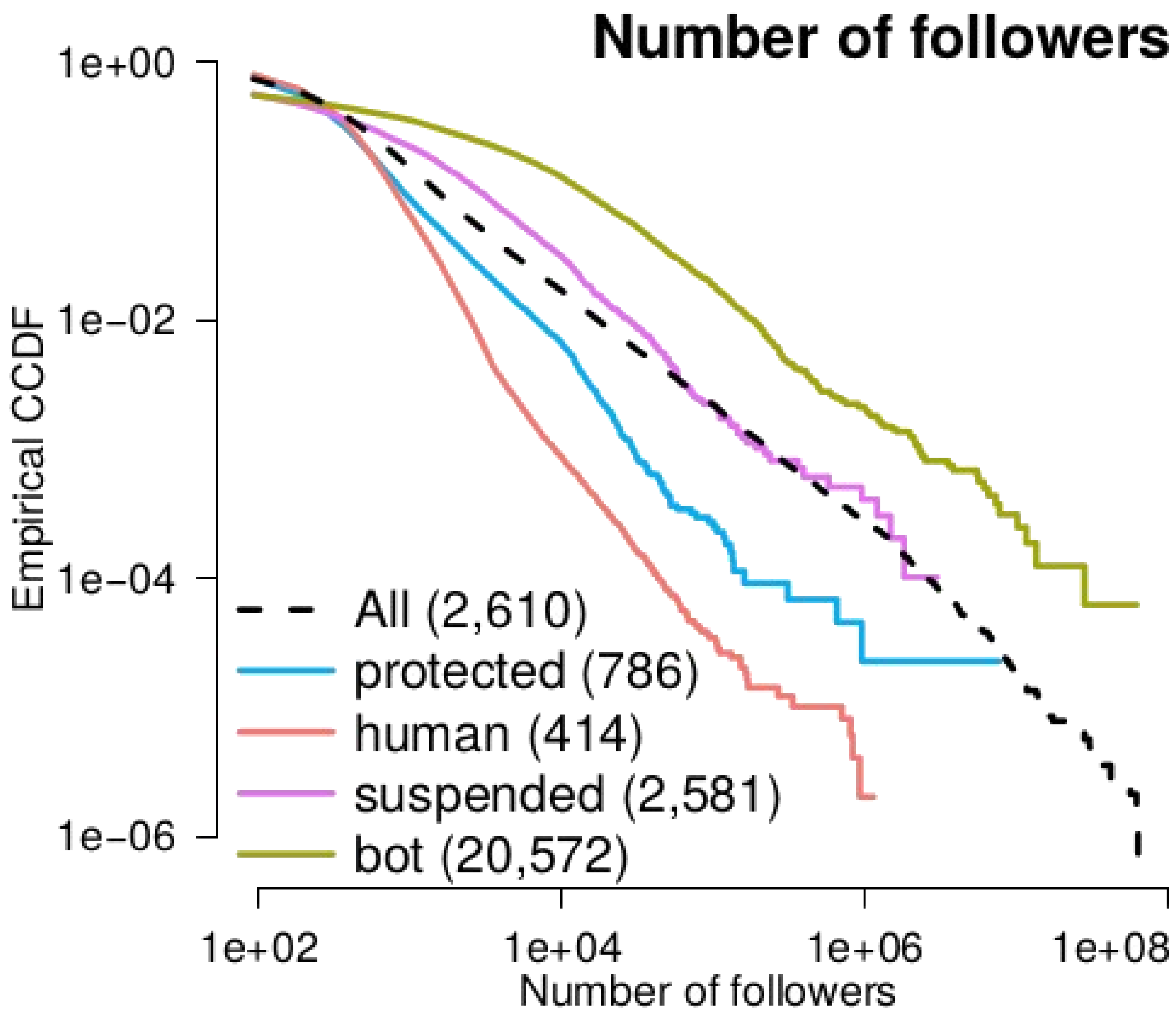
User botness



Analyze political behavior of bots

Activity profiling

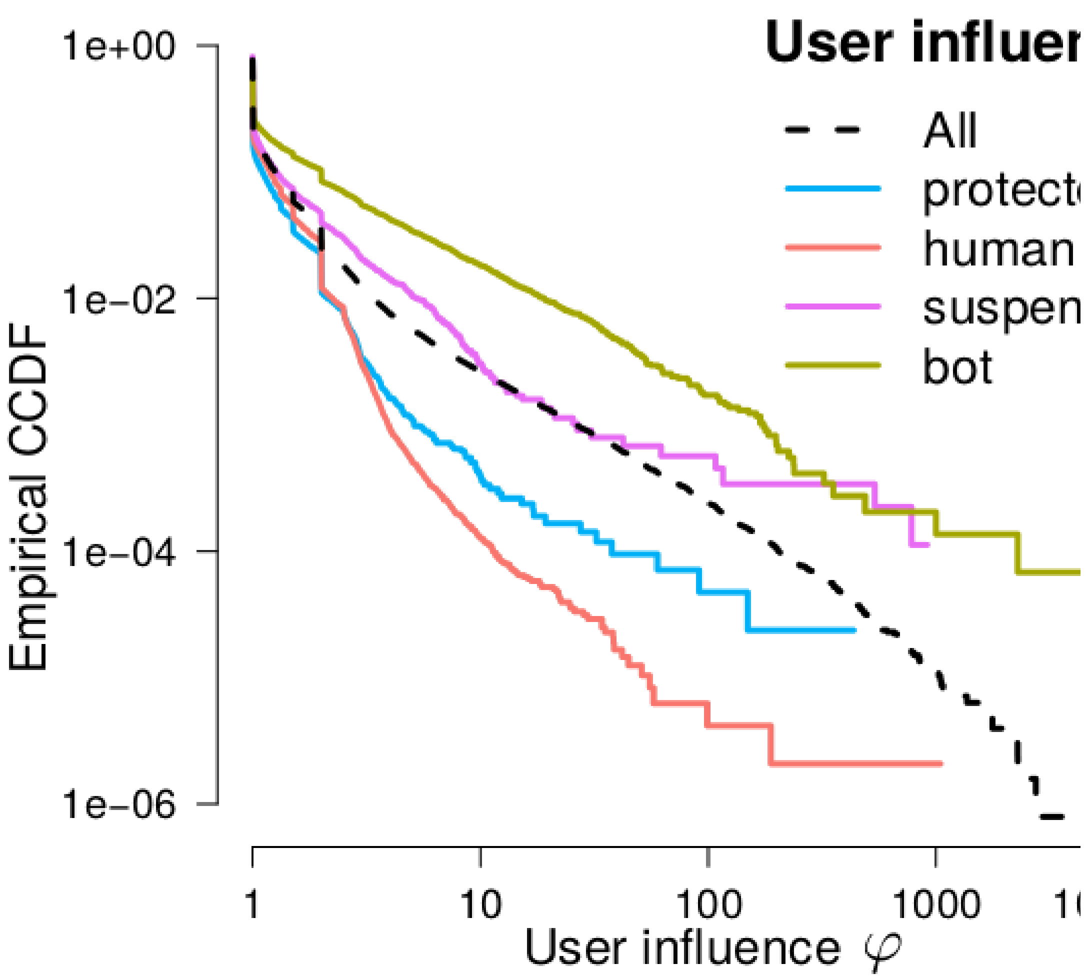


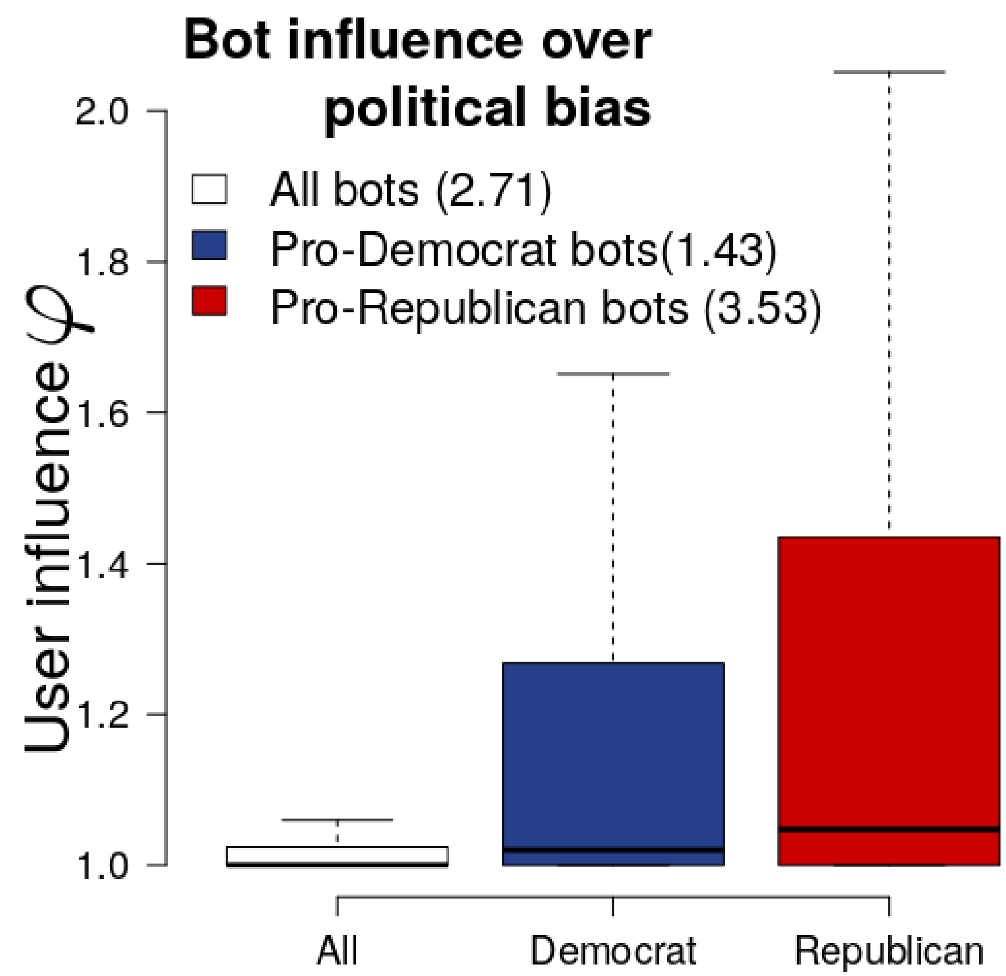


Bots and **Suspended** are more active than **Humans** and **Protected**

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

User influence

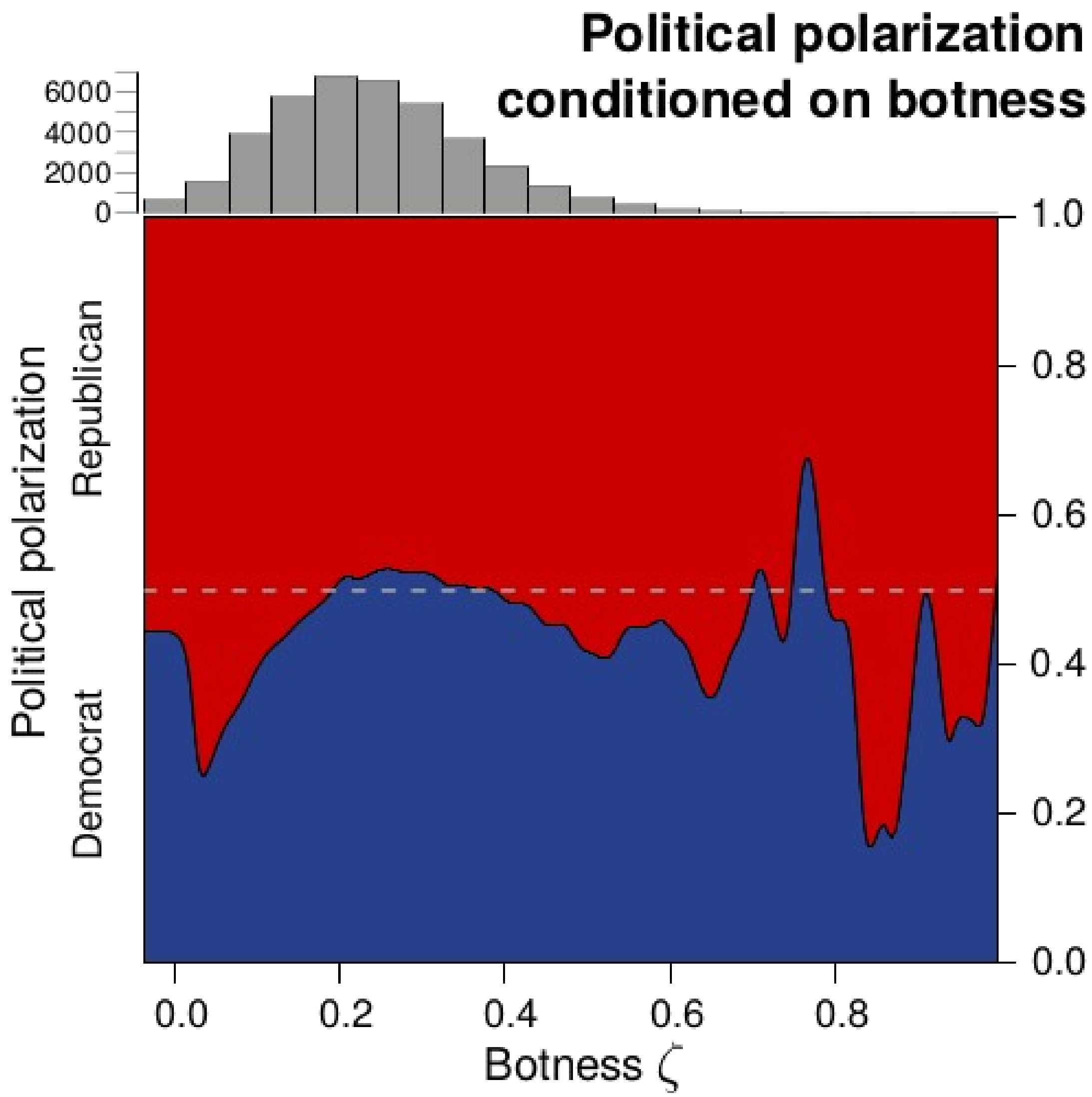


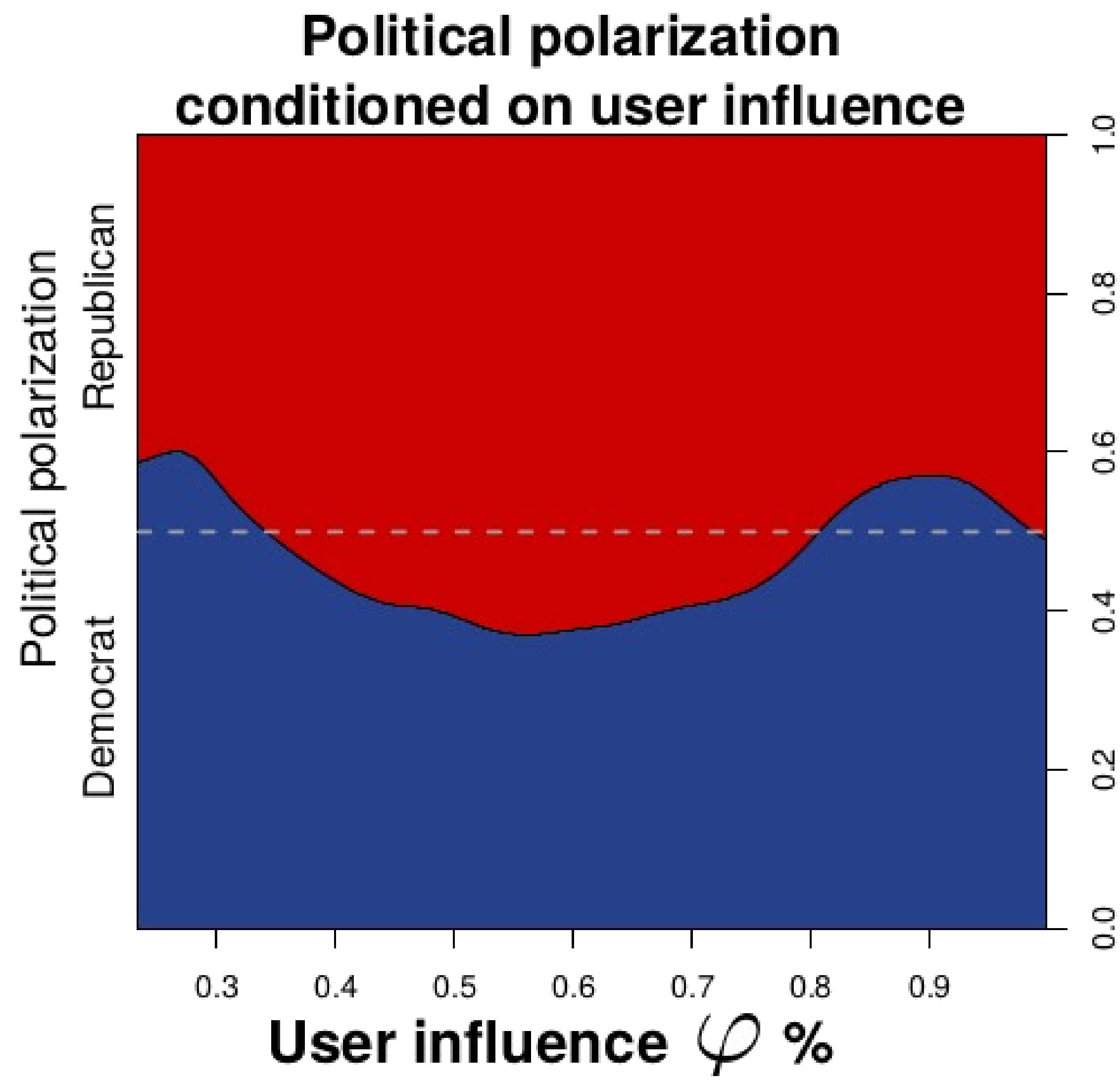


The average **Bot** has 2.5 times more influence than the average **Human**

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

Political partisanship

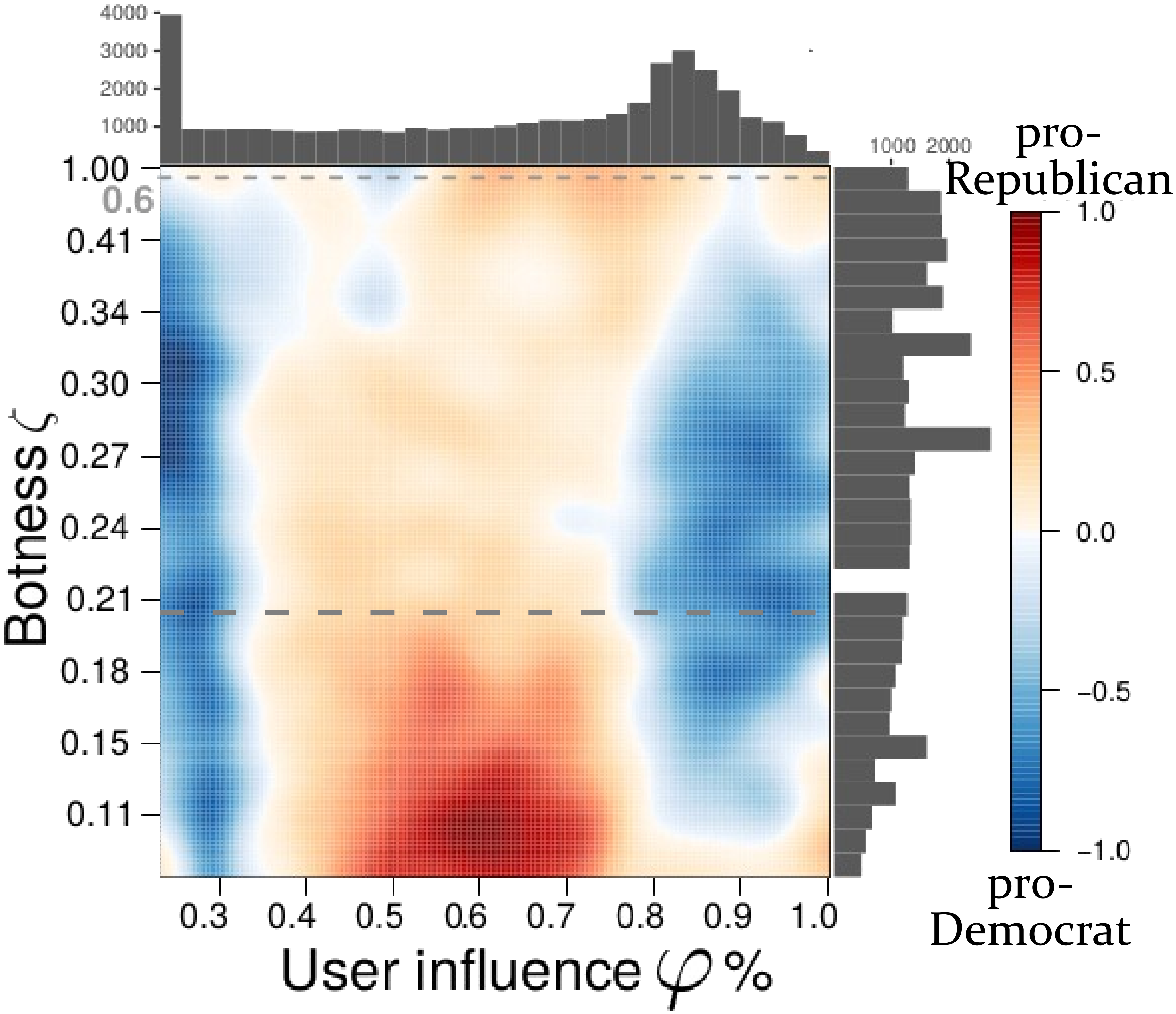




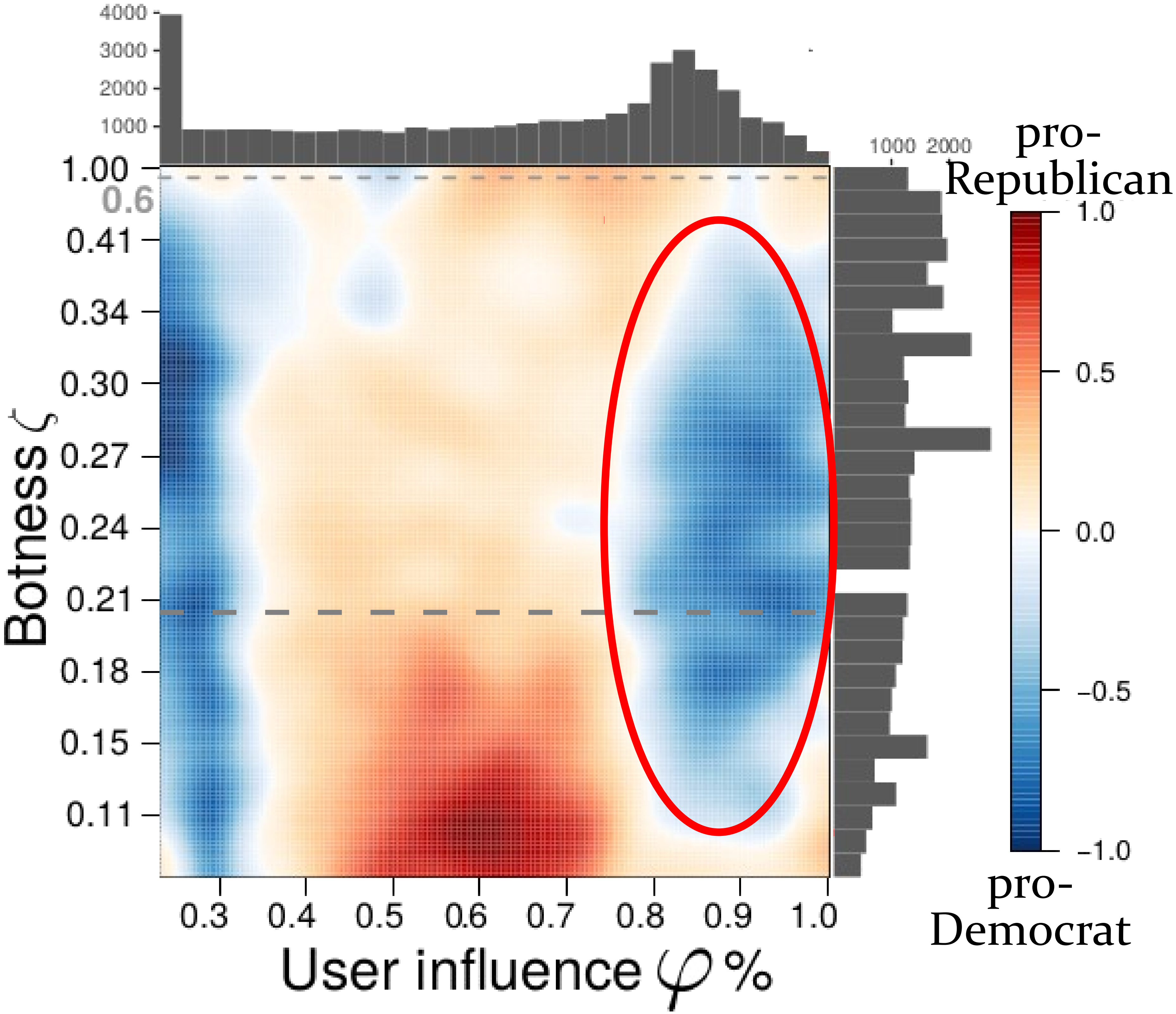
Bots are more likely to be pro-Republican (than pro-Democrat)

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

Polarization map

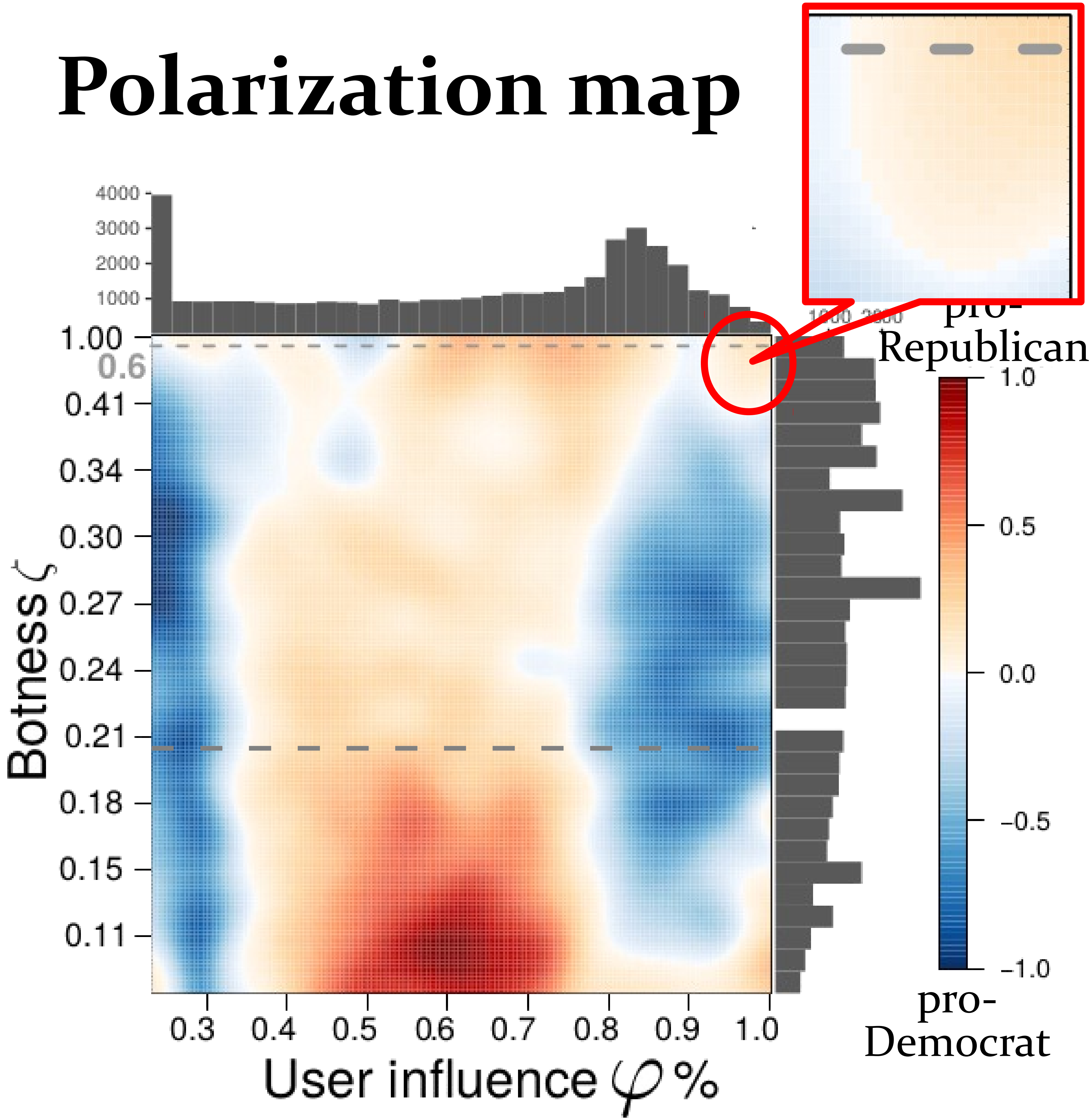


Polarization map



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

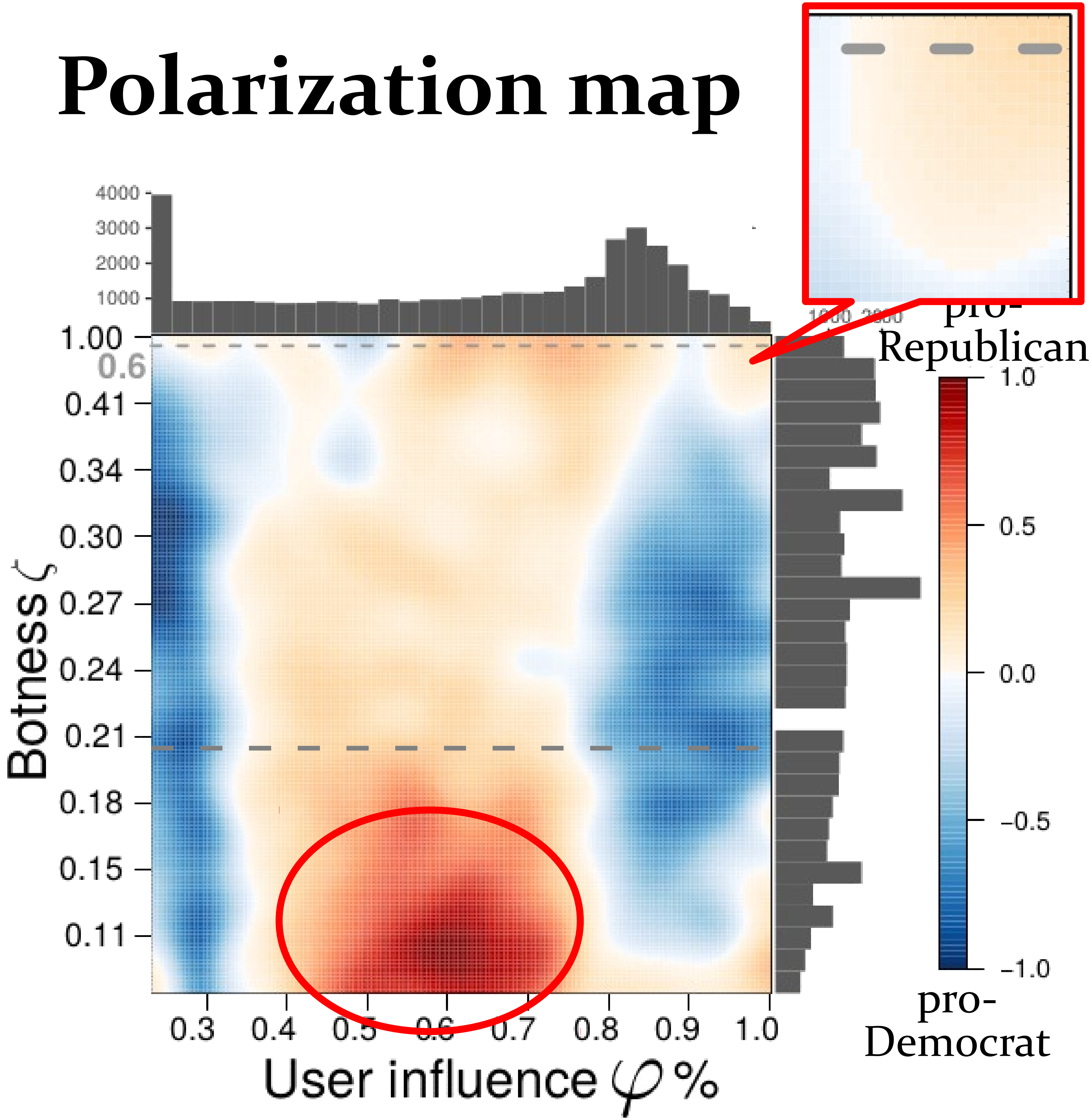
Polarization map



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

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

Polarization map

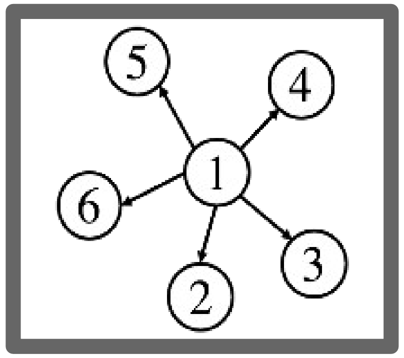


Very highly influential users are pro-Democrat (**D: 7201**, **R: 5736**)

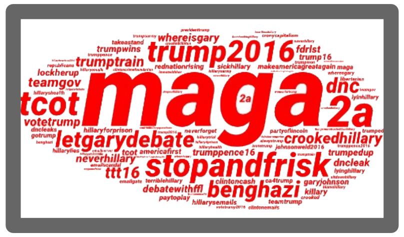
Highly influential **Bots** are pro-Republican (**D: 24**, **R: 45**)

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

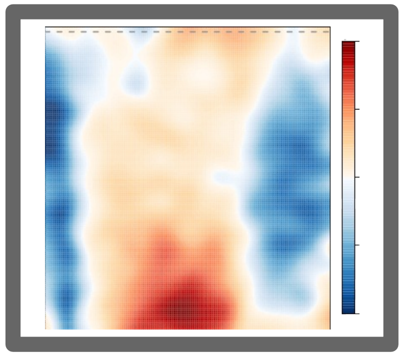
Summary



A scalable algorithm to estimate user influence from latent network structures

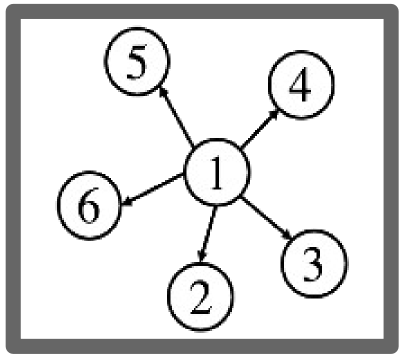


Three measures to quantify the influence, the political partisanship and botness of Twitter users



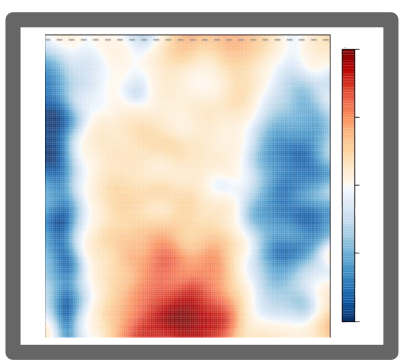
A detailed analysis of the role and influence of socialbots during the first U.S. Presidential debate.

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



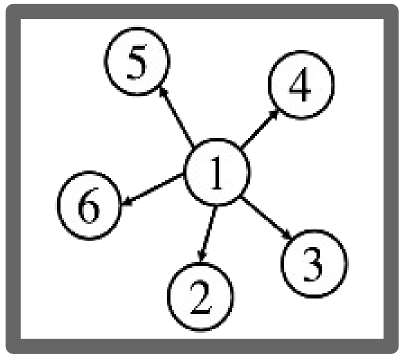
A detailed analysis of the role and influence of socialbots during the first U.S. Presidential debate.

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

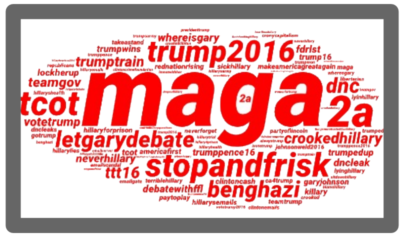
Were Bots instrumental for the results of the elections?

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

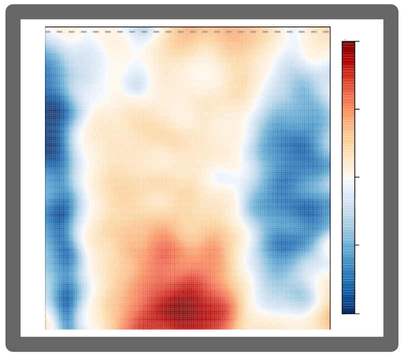
https://github.com/computationalmedia/cascade-influence



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.

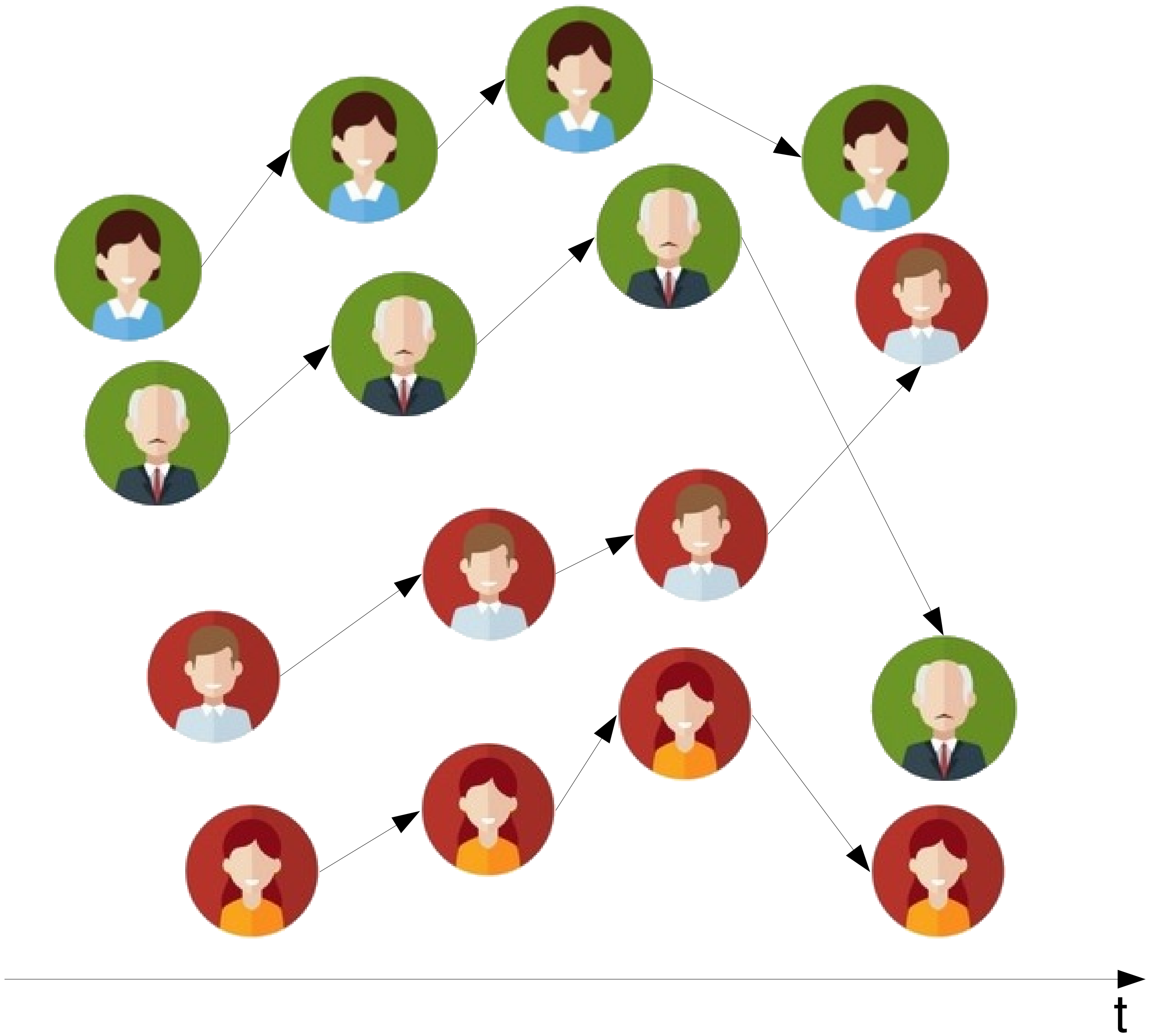
Limitations: Organizational accounts appear as **Bots**; binary partisanship characterization (e.g. independent voters)

Were Bots instrumental for the results of the elections?

User identity via semantic edit distance: A case study of Russian trolls on Twitter

[Kim et al Jour. Comp. Social Science '19]

Behavioral Data Science



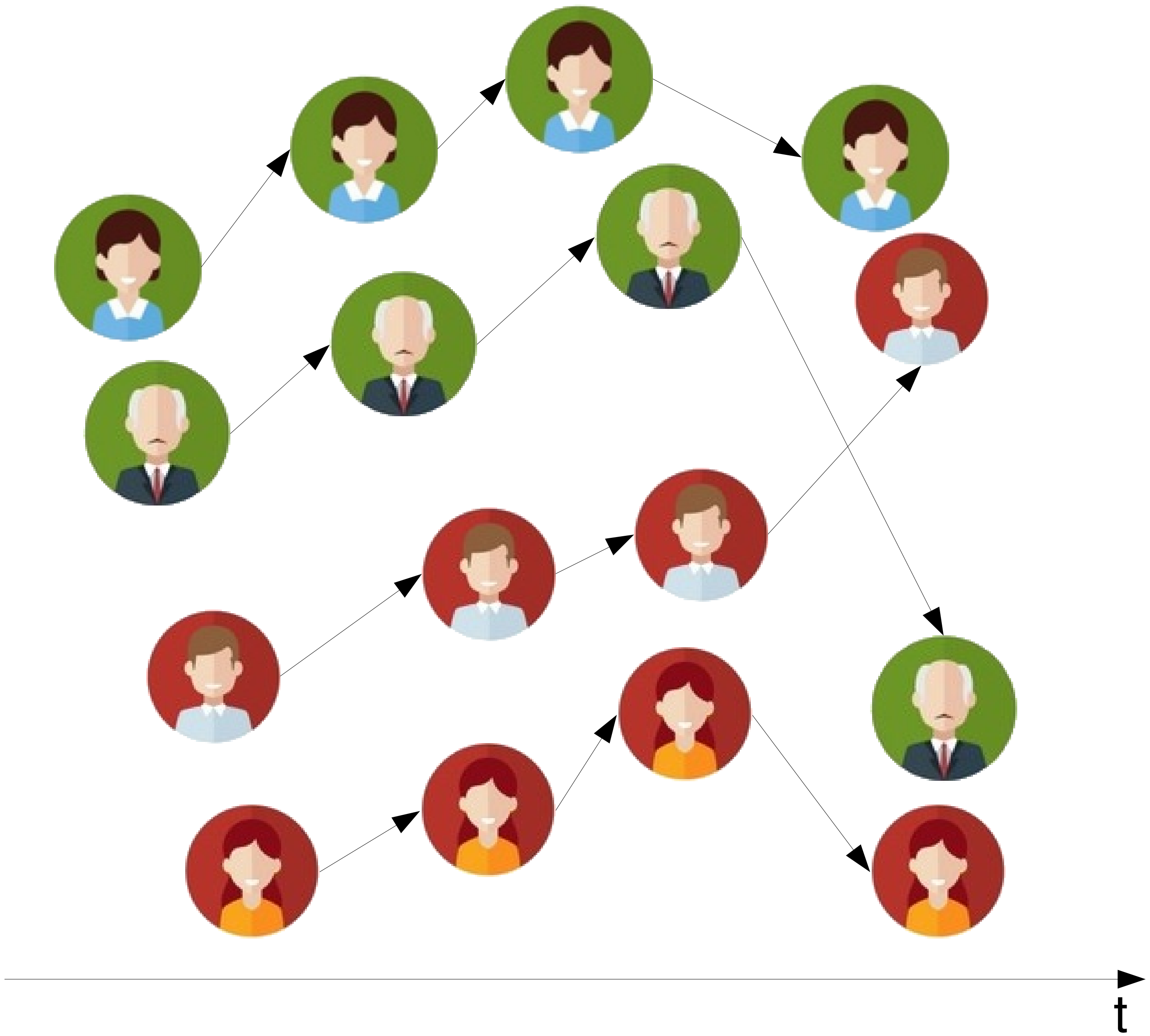
Identity through the digital traces that actors leave behind

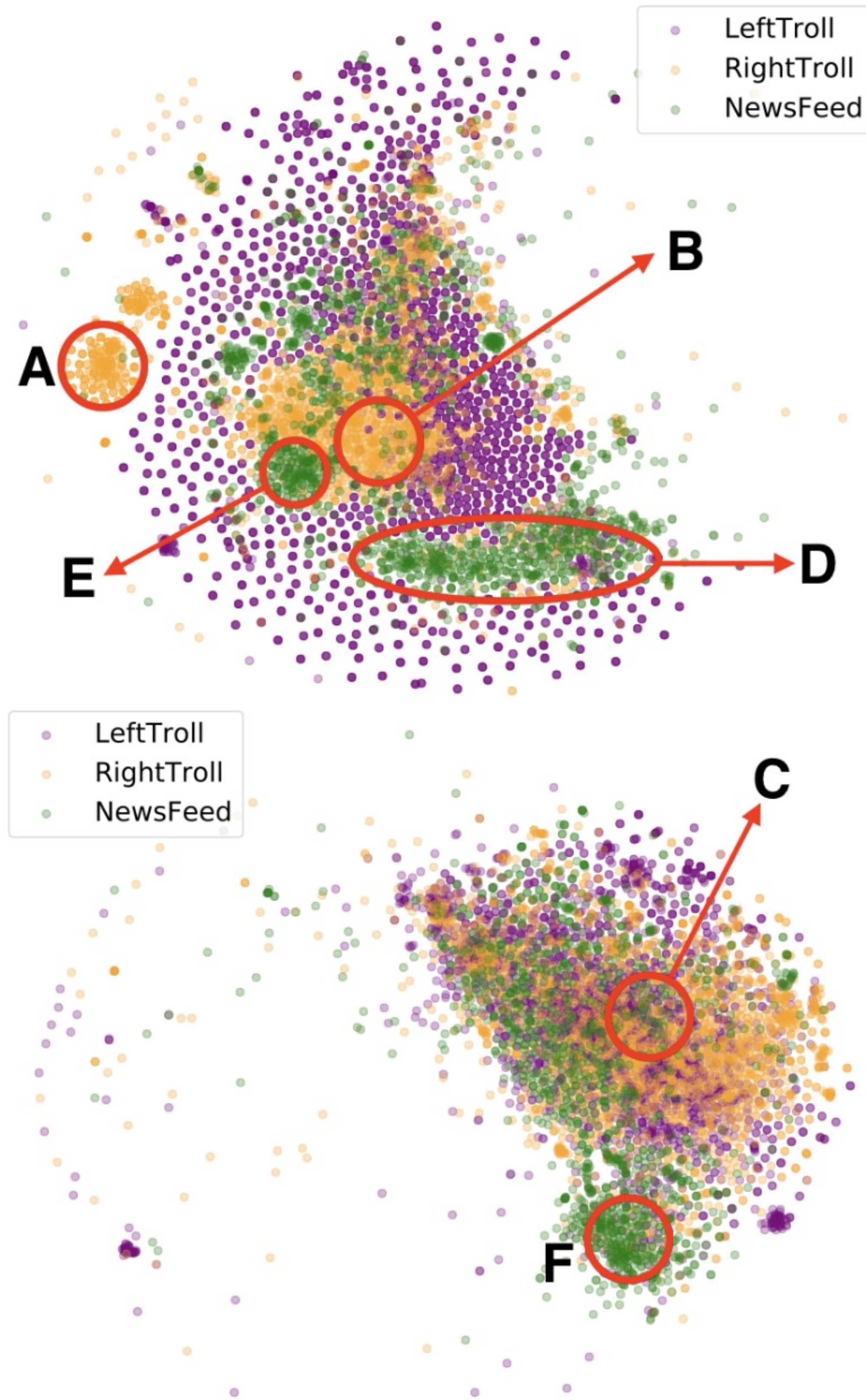
Kim, D., Graham, T., Wan, Z., & Rizoiu, M.-A. (2019). Analysing user identity via time-sensitive semantic edit distance (t-SED): A case study of Russian trolls on Twitter. Journal of Computational Social Science (under review). https://arxiv.org/abs/1901.05228

User identity via semantic edit distance: A case study of Russian trolls on Twitter

Behavioral Data Science

[Kim et al Jour. Comp. Social Science '19]

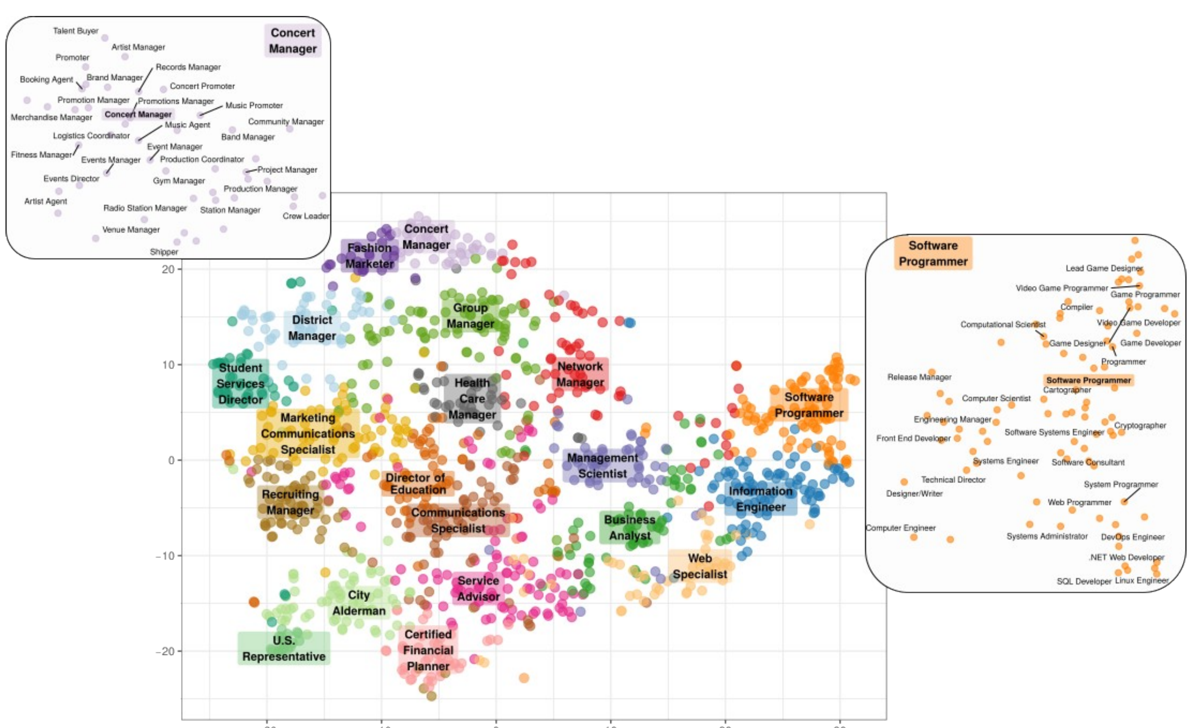




Identity through the digital traces that actors leave behind

Kim, D., Graham, T., Wan, Z., & Rizoiu, M.-A. (2019). Analysing user identity via time-sensitive semantic edit distance (t-SED): A case study of Russian trolls on Twitter. Journal of Computational Social Science (under review). https://arxiv.org/abs/1901.05228

Thank you!



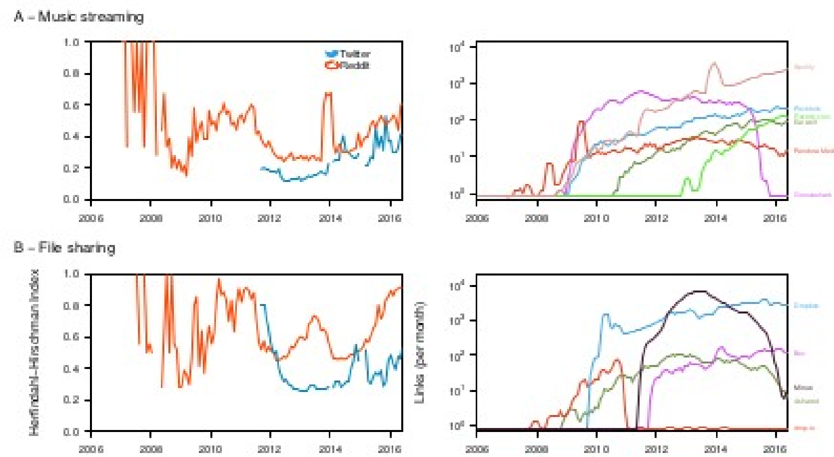
Other projects

Other projects

Behavioral Data Science

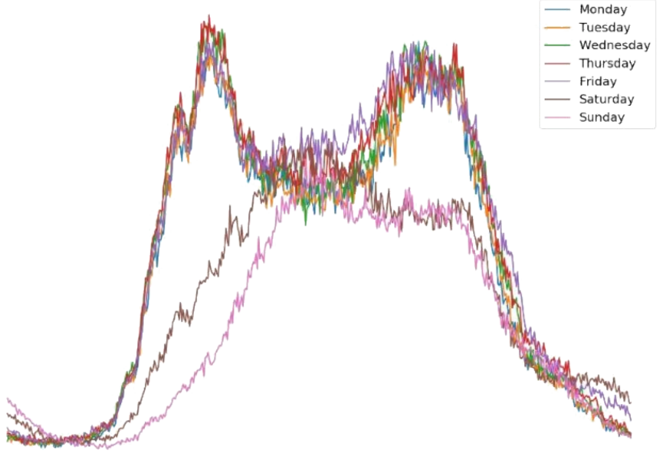
Wikipedia privacy

[Rizoiu et al WSDM‘16]



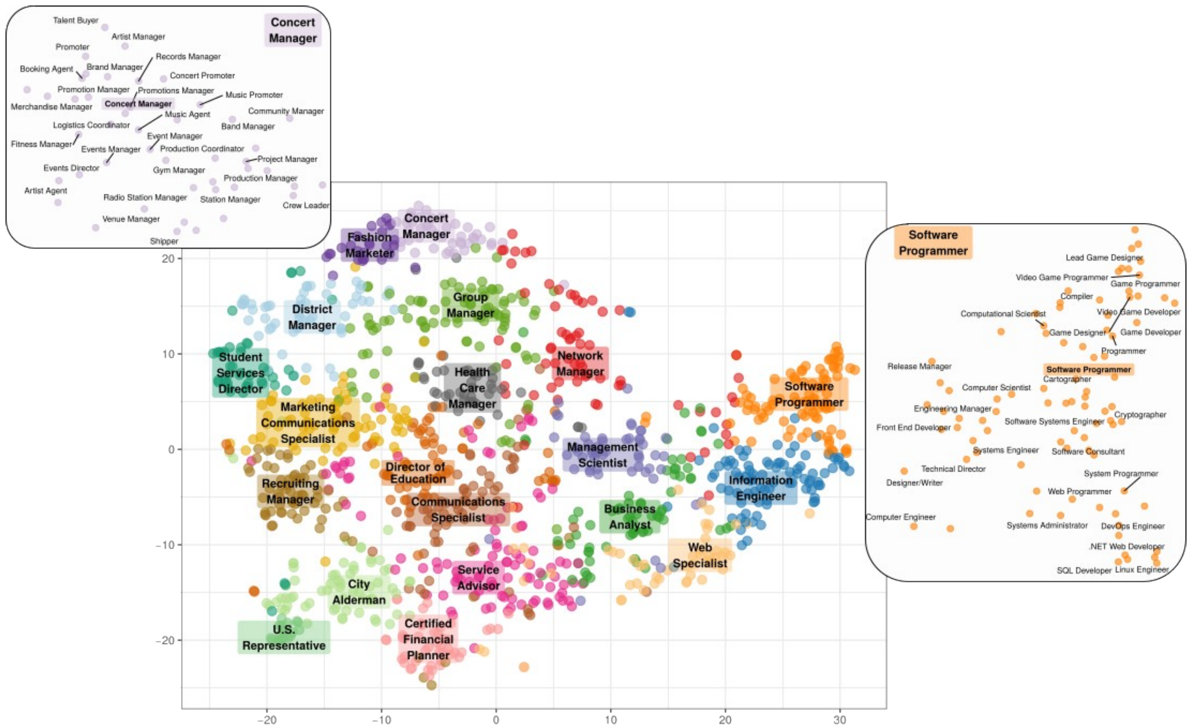
Online Diversity

[McCarthy et al ‘19]



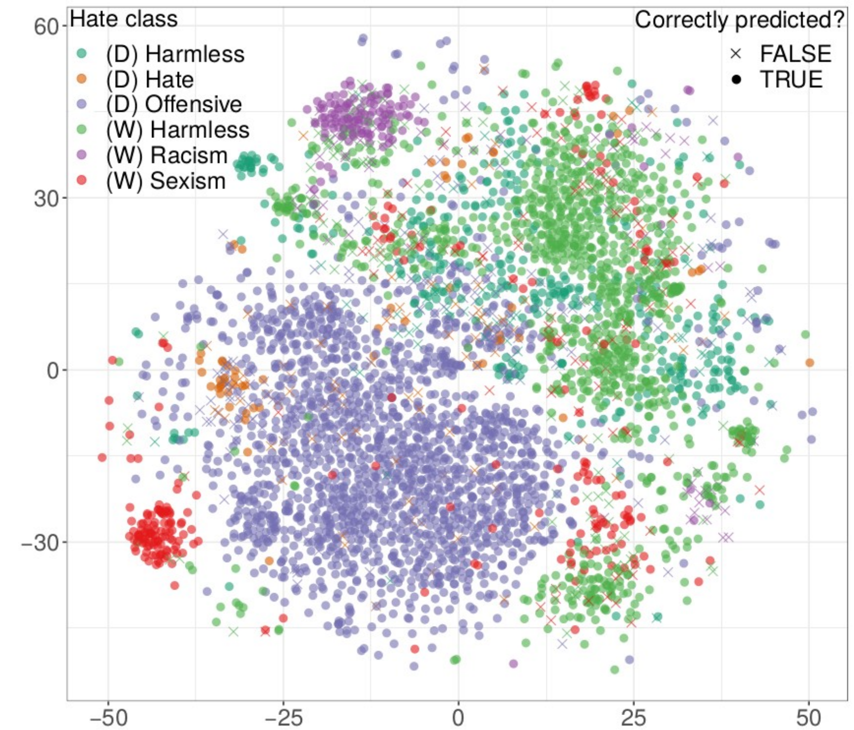
Smart traffic

[Mihaita et al ITSC‘19]



Vocation compass

[Kern et al PNAS‘19]



Transfer learning for Hate Speech detection

[Rizoiu et al ICWSM‘19]

Other projects – references

[Rizoiu et al WSDM'16] Rizoiu, M.-A., Xie, L., Caetano, T., & Cebrian, M. (2016). Evolution of Privacy Loss in Wikipedia. In International Conference on Web Search and Data Mining (WSDM '16) (pp. 215–224). New York, New York, USA: ACM Press. http://arxiv.org/pdf/1512.03523.pdf

[McCarthy et al '19] McCarthy, P. X., Rizoiu, M.-A., Eghbal, S., & Falster, D. S. (2019). Long-term evolutionary trends of diversity online.

[Mihaita et al ITSC'19] Mihaita, A.-S., Li, H., He, Z., & Rizoiu, M.-A. (2019). Motorway Traffic Flow Prediction using Advanced Deep Learning. In 22nd Intelligent Transportation Systems Conference (ITSC'19).

[Kern et al PNAS'19] Kern, M. L., McCarthy, P. X., Chakrabarty, D., & Rizoiu, M.-A. (2019). Social Media-Predicted Personality Traits Can Help Match People to their Ideal Jobs. Proceedings of the National Academy of Sciences (under review).

[Rizoiu et al ICWSM'19] Rizoiu, M.-A., Wang, T., Ferraro, G., & Suominen, H. (2019). Transfer Learning for Hate Speech Detection in Social Media. International AAAI Conference on Web and Social Media (ICWSM'19) (under review). http://arxiv.org/abs/1906.03829