MAPPING THE ECHO-CHAMBER: DETECTING AND CHARACTERIZING PARTISAN NETWORKS ON TWITTER.

THOMSON REUTERS

{Armineh Nourbakhsh, Xiaomo Liu, Quanzhi Li, Sameena Shah}@thomsonreuters.com

Research and Development, Thomson Reuters

RESEARCH & DEVELOPMENT

BACKGROUND & PROBLEM STMT

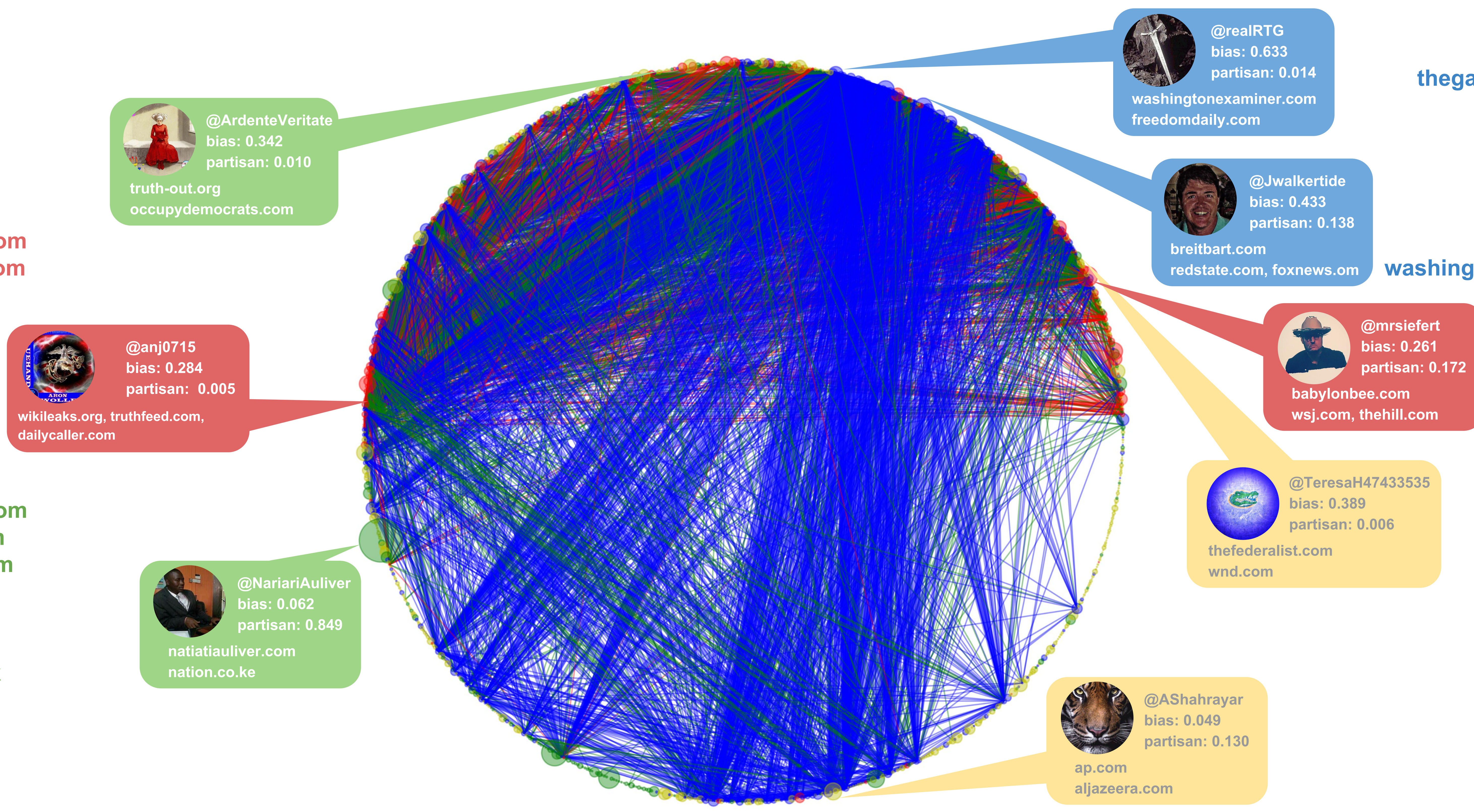
- Echo chambers: communities that share content with similar viewpoint; rarely interact with other communities in positive terms
- Studies that have focused on echo chambers:
  - Use real network of follower/friends, but:
    - Complete network info is hard to obtain.
    - Follower/followee relations don't necessarily reflect political agreement.
  - Use labeled data, but:
    - Labels may be biased themselves.
    - Labels may change over time, e.g. many former UKIP supporters voted for Labour in UK's generation election of '17.
  - After echo-chambers are identified, all of their content is dismissed as biased.
- **Problem**:
  - Identify echo-chambers on Twitter, such that:
    - ***Network info*** is not required.
    - No ***human labeling*** is required.
    - Model can identify ***which topics are subject to misrepresentation*** by each echo-chamber.
  - Do all of this in a way that the model can be updated in ***real-time***.

DATA & METHODOLOGY

- Seed set: domains mentioned in Kaggle's fake news dataset (2016-10-25 to 2016-11-25).
- Download all tweets that mentioned those domains within the same time period: **25,742,785** tweets from **131,552** unique accounts
- Use **PMI**-like measure to score domains by
  - similarity to each other
  - similarity to Kaggle domains
- Score users based on the domains they share:
  - **Bias**: similarity of a given user's domains to Kaggle domains.
  - **Partisanship**: similarity of a given user's domains to each other.
- Draw network of co-shared domains:
  - **Node size**: partisanship of each user.
  - **Edge width**: number and cumulative bias of two nodes' shared domains.
- Run **Louvain modularity detection** to identify echo-chambers (color-coded).
- Calculate each chamber's deviation:
  - **Word2Vec deviation** from a large global set.
  - **TF-IDF deviation** from other chambers.
- **High Word2Vec deviation AND low TF-IDF deviation ⇒ potentially misrepresented topic.**

usatoday.com
nbcnews.to
nymag.com
theonion.com
dailykos.com
theintercept.com
businessinsider.com
insider.foxnews.com
reddit.com
economist.com

washingtonpost.com
newyorktimes.com
huffingtonpost.com
cnn.com
theguardian.com
change.org
independent.co.uk
thehill.com
bbc.com
latimes.com



thegatewaypundit.com
foxnews.com
breitbart.com
dailycaller.com
truthfeed.com
dailymail.com
infowars.com
washingtonexaminer.com
zerohedge.com
nypost.com

pscp.tv
buzzfeed.com
medium.com
time.com
wsj.com
forbes.com
amazon.com
npr.org
telegraph.co.uk
bloomberg.com

FBI, bad, witch | covering, Mueller, investigated | Obama, Comey, million | shooter, apology, MSM

death, travel, Muslims | wikileaks, alleged, brothers | vaccines, reckless, study

Tory, legal, politician | Corbyn, #LastMinuteCorbynSmears, @jeremycorbyn | #bbcsp, industry, MSM

EVALUATION

- Buzzfeed's list of **60 true stories and 60 fake stories** propagated on Facebook in the period leading up to 2016 U.S. presidential election.
- Download tweets linking to the same articles (**no seed set ⇒ no bias score**).
- Apply echo-chamber analysis to the users.
- Train linear classifier based on: **user partisanship, domain exclusivity, content match against each chamber's set of deviating terms**
- F1 scores show the features are robust & informative, even for simple classifiers

Features	LogReg	DTree	SVMLinear
user + domain	0.73	0.71	0.74
user + content	0.69	0.67	0.69
domain + content	0.75	0.70	0.76
All	**0.77**	**0.73**	**0.78**

DISCUSSION

- Suggested setup for real-time setting:
  - Network of chambers (re)constructed on a daily/weekly basis
  - An index constructed to score domains based on their popularity in each chamber
  - Tweets stream from the public API (e.g. using Spark streaming)
  - Stories are tracked by their URL
  - Users are benchmarked against the network
    - To avoid exhausting rate limits, a KB of users' common domains can be stored in a cache
  - Each story's score based on:
    - **User scores**
    - **Content score**
    - **Domain score**

SEE WHERE YOU STAND

TRY THE LIVE DEMO