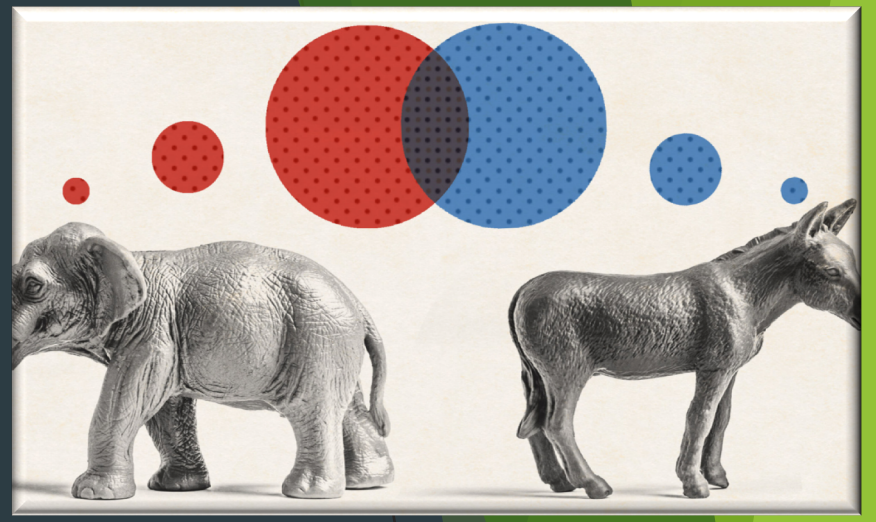
The Polarization of Information

Professor Narayana Prasad Santhanam

Charles Dickens

Dept. of Electrical Engineering, University of Hawaii at Manoa

Introduction and Motivation



TED (2016)

- Many would agree that the United states political climate is becoming increasingly polarized
  - How could we quantify this troubling intuition?

- Is left vs. right enough to capture what is truly happening on the web?
- How does polarization vary across topics?

Project Description

- The primary focus of this project is to develop a system that can provide objective measurements of the polarity of the network of information and opinions on the popular microblog, Twitter.

- The short term objectives of this project are:
  1. Develop a method for comparing the polarization of existing communities of thought at the topic specific level
  2. Improve methods for modeling social networks

- The contributions of this report is the framework for a workflow which:
  1. Models a network of tweets related to a common topic using a semi-supervised approach
  2. Finds community structure in the model
  3. Measures the polarity of the network using graph conductance

Wired (2018)

Methods

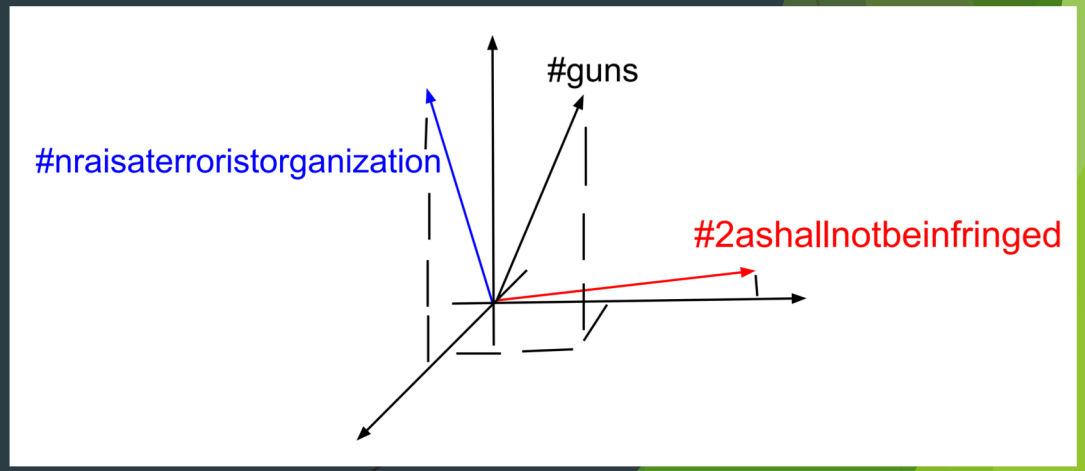
- Retrieving Relevant Information and Opinions:
- Twitter granted this project access the premium search API sandbox, which allows for 25000 tweets to be collected each month.

Topic	Queries
Gun Control	'guns'
Voting	'voting rights', 'voter turnout', 'voter fraud' 'voter suppression', 'vote'
Immigration	'immigration', 'illegal immigrants', 'trump wall' 'border security', 'asylum seekers'

Table 1: Twitter API Search Queries by Topic

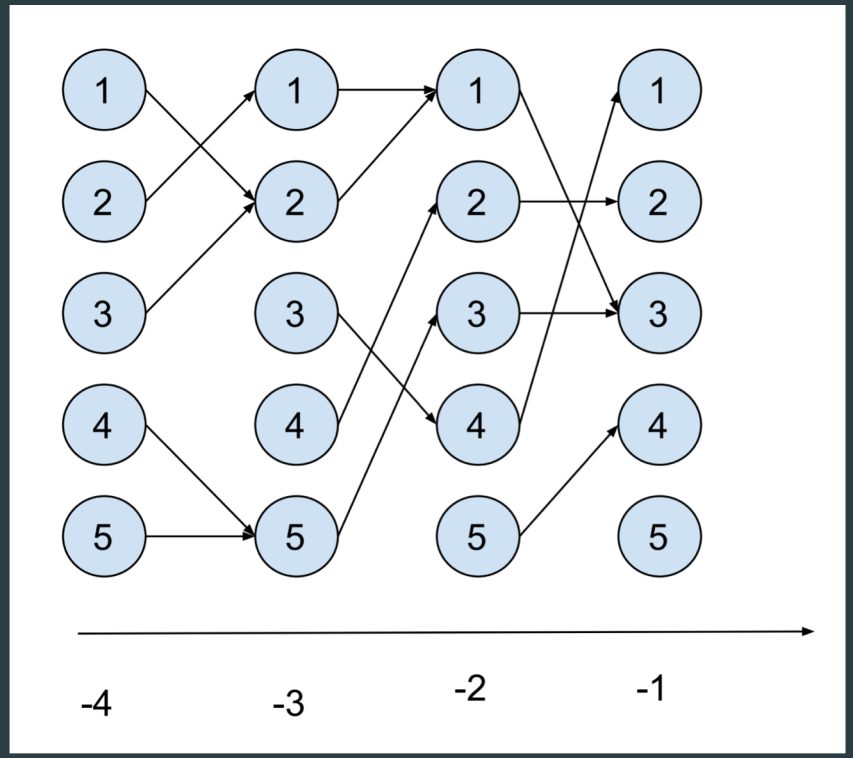
Methods

- Modeling the Network:
- The set of all unique hashtags, H, is aggregated into a file and is manually labeled with a sentiment value existing in the set {-1, -1/2, 0, 1/2, 1}
- Tweets, $t_i \in T$, are modeled as nodes in the network and an edge between two nodes is weighted to reflect how similar the two tweets are.
- Edge weights in the network are learned using a semi-supervised approach that ensures Tweets with highly contrasting hashtag sentiment are not clustered together.



Methods

- Community Detection:
- The community detection algorithm used in the system, backward path community detection, was developed at the University of Hawai'i Big Data Lab (Paravi and Santhanam, 2015)



Methods

- Polarity Calculations
- The polarity is measured using an estimate of the graph conductance metric.
- The Conductance of a cut (S, S) of a graph $G = (V, E)$ is defined as

$$\varphi(S) = \frac{\sum_{i \in S, j \in \bar{S}} w_{i_j}}{\min(w(S), w(\bar{S}))}$$

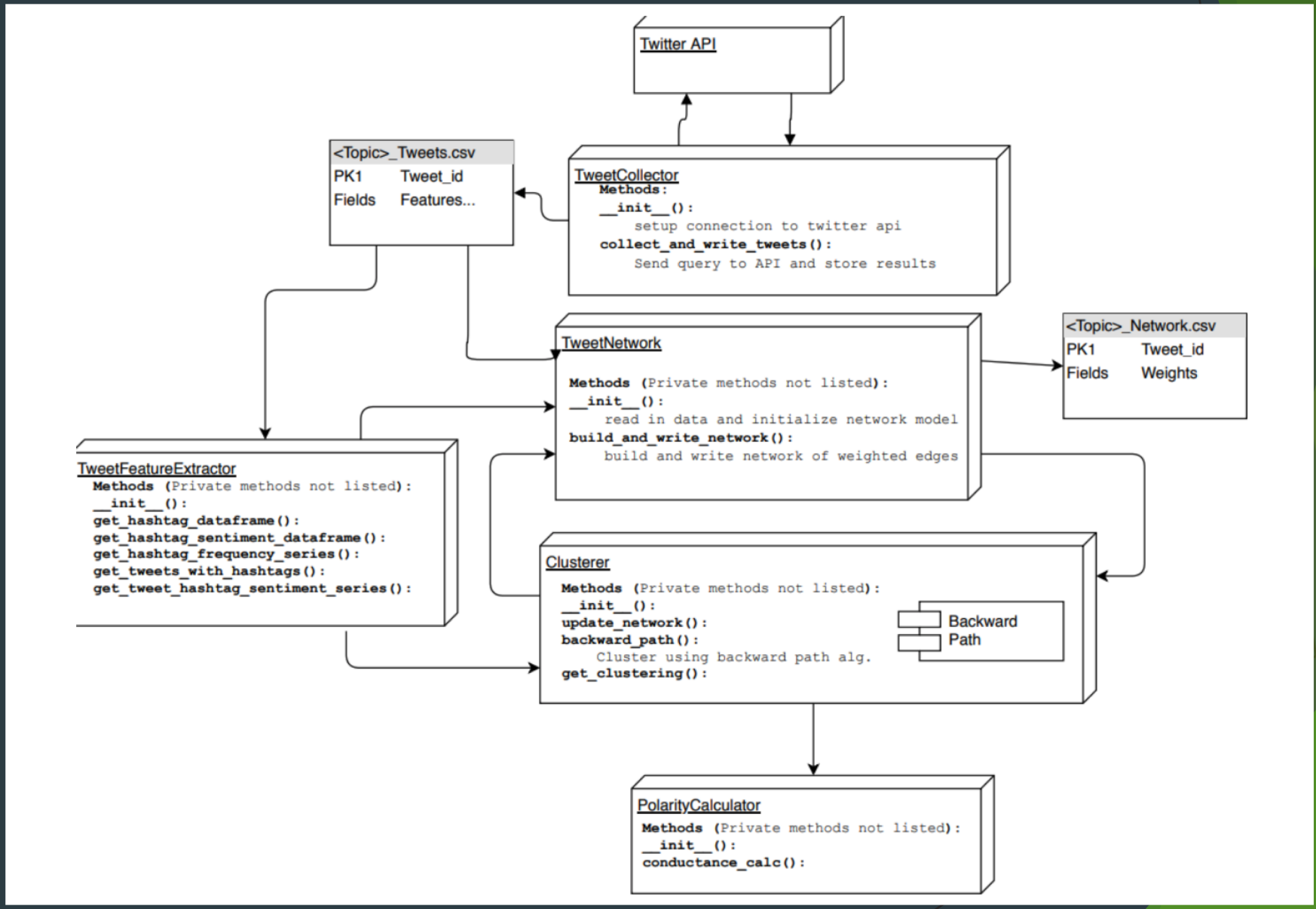
where

$$w(S) = \sum_{i \in S} \sum_{j \in V} w_{i,j}$$

Then the graph conductance is defined as

$$\phi(G) = \min_{S \subseteq V} \varphi(S).$$

Implementation



Results

Topics	Count of Tweets w/ Hashtags	Count of Unique Hashtags
Gun Control	500	408
Voting	847	401
Immigration	356	187

Table 2: Twitter API Search Results by Topic

Results

Cluster	text
0	Lets get real on homelessness & ... #bbcqt
0	"1 in 200" are homeless,.. #bbcqt ... Blame endless immigration...
0	... taxpayer funded benefits and homes go to immigrants... bbcqt
1	#Thanksgiving.... Native-Americans understood... immigration are our strengths...
1	#Thanksgiving the commemoration of a day in which 50 Americans saved the lives of 52 illegal immigrants from England.
2	Mr. President we are some fed up Americans. We want to #BuildTheWall and be done with illegal immigrants.
2	#BuildTheWall Trump Super PAC Calls on Congress to Fund the Wall via BreitbartNews
2	PRESIDENT TRUMP... Border Wall Funding #BuildTheWall #CaravanInvasion #RWeBComingEurope
2	Deal with Mexico paves way for asylum overhaul at U.S. border @realDonaldTrump #BuildTheWall #TheGreatAwakening #QA

Table 4: Sampling of tweet texts from different communities identified from the 'immigration' data set

Results

Network model visualizations generated using Fruchterman-Reingold force-directed algorithm

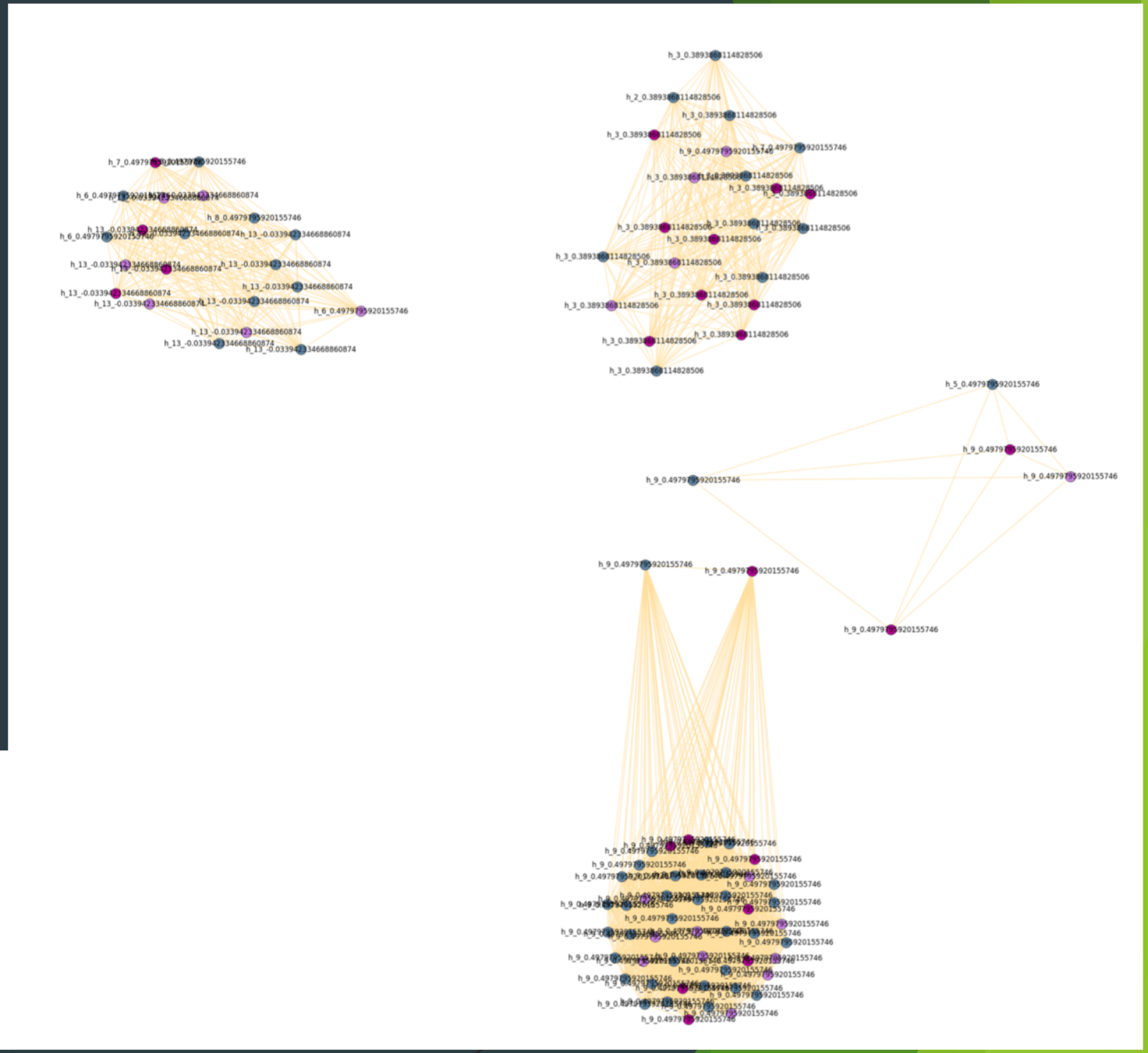
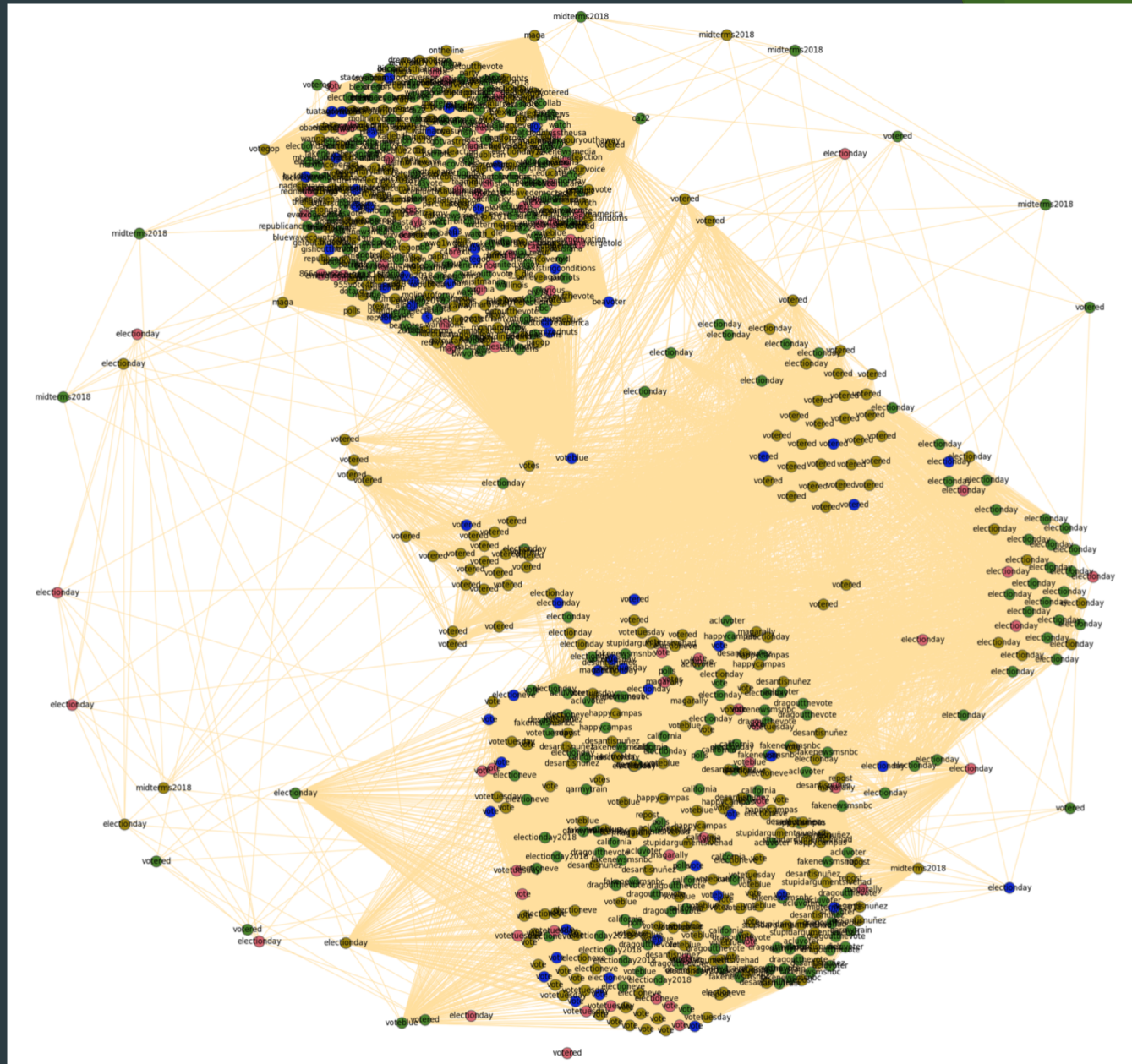


Figure 1: Simulated network using LFR benchmark and hashtag distribution based on bernoulli trials with probabilities dependent on the tweet's true community assignment.

Results

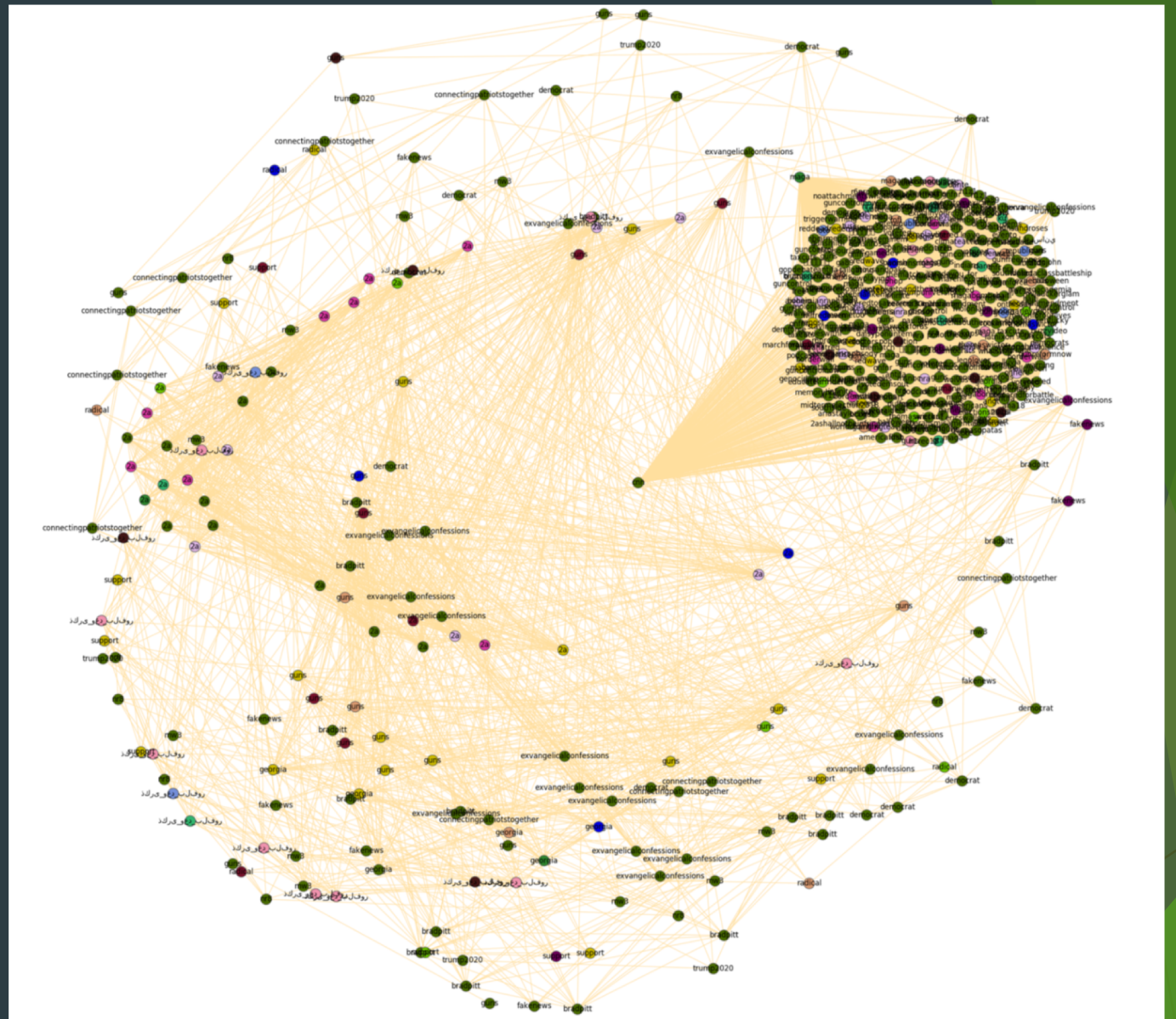
Network model visualizations generated using Fruchterman-Reingold force-directed algorithm



(a) Voting Rights and Turnout

Results

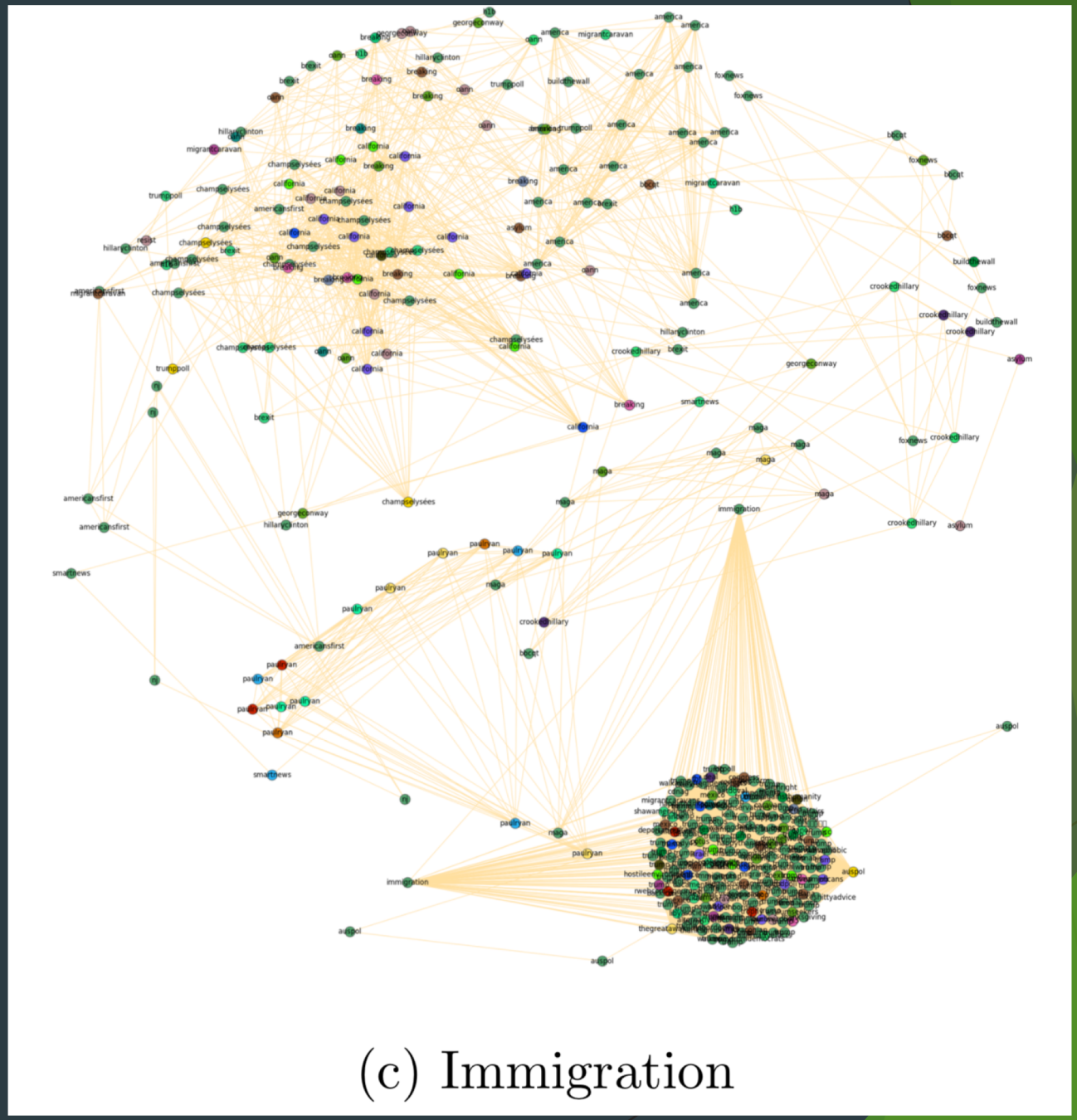
Network model visualizations generated using Fruchterman-Reingold force-directed algorithm



(b) Gun Policies

Results

Network model visualizations generated using Fruchterman-Reingold force-directed algorithm



(c) Immigration

Conclusions

- This implementation has been a proof of concept for the techniques used to build the network model, identify communities, and measure polarity.
- The system will be developed in future work to consider additional features such as tweet popularity, tweet text sentiment, and more.

Acknowledgements

Special thanks to

CSoI: Center for Science of Information (https://www.soihub.org) is a National Science Foundation Science and Technology Center made possible under grant NSF CCF- 0939370

References

- TED (2016). *Political Polarization Illustration*. [image] Available at: https://tedideas.files.wordpress.com/2016/11/featured_art.jpg?w=750 [Accessed 4 Dec. 2018].
- Wired (2018). *Twitter-FeatureArt*. [image] Available at: https://media.wired.com/photos/5941a9186600f013e9634520/191:100/pass/Twitter-FeatureArt.jpg [Accessed 5 Dec. 2018].