

# Relating Blog Political Bias to Network Structure

Lucien Carroll

2006 May 16

Ling 795: Web Text

# Procedure

- Collected pages from each site, recording all URLs seen (HarvestMan)
- Based on first 500 URLs from each site, constructed network graph
- Use information centrality to identify cores of the two major camps
- Assign other nodes to camps according to distance to the cores
- Correlate centrality and isolation with bias

# Motivation

- It is often very helpful to know “where the writer is coming from”
  - Searching “partial-birth abortion” returns mostly pro-life pages
  - Searching “San Salvador Atenco” turns up mostly pro-Zapatista pages
- Beyond prestige or popularity, we want to
  - Recognize coherent ideological camps
  - Identify which camp a page belongs to
  - Identify a page's degree of bias

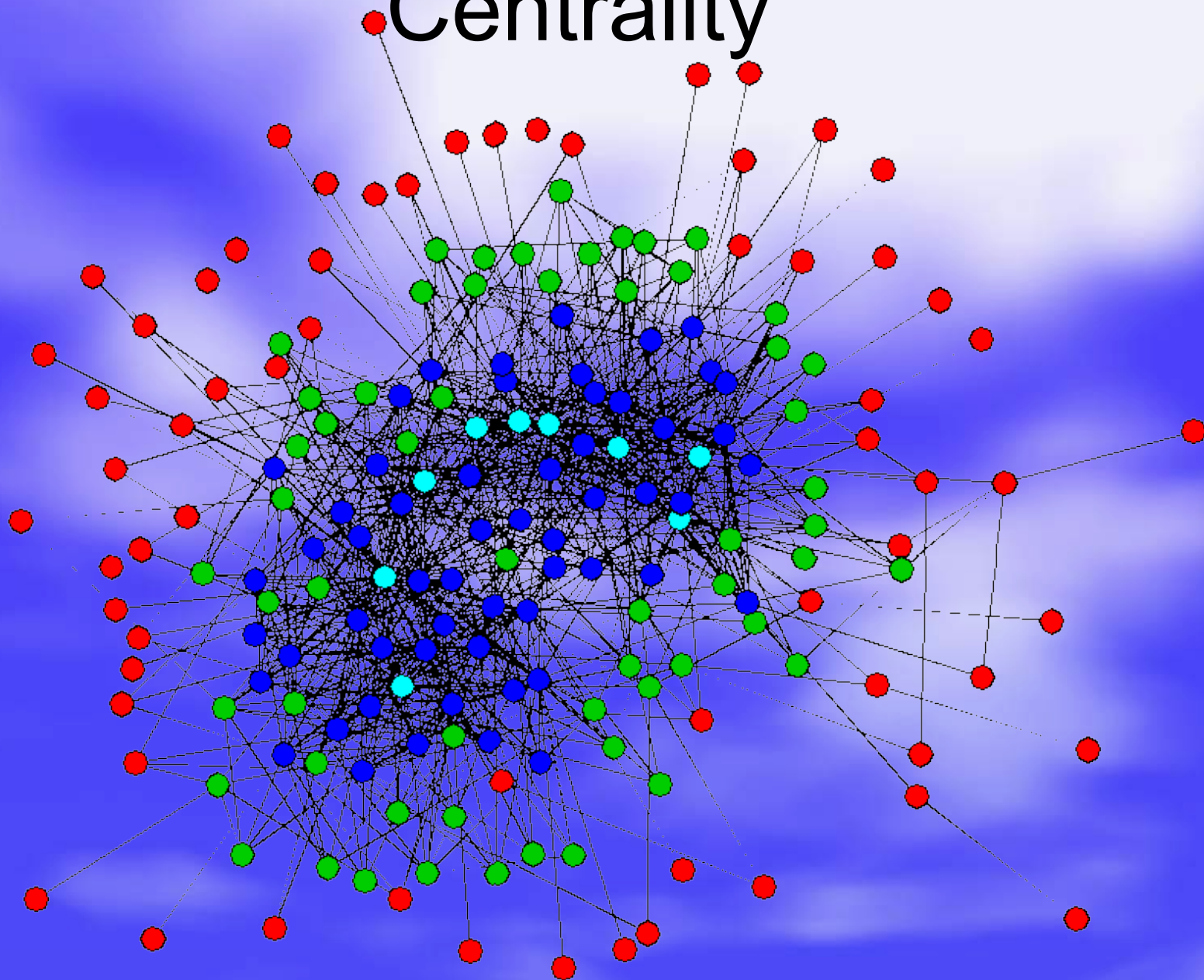
# Background

- We know ideological camps are revealed in network structure
  - Books on Amazon.com, Debate about European Constitution, Political websites in general, Political blogs, Political discussion boards
  - But topical camps have similar network structures
    - Flake et al's virtual communities, Musicians' websites
- This has been concerned with identifying membership, when given ideological camps

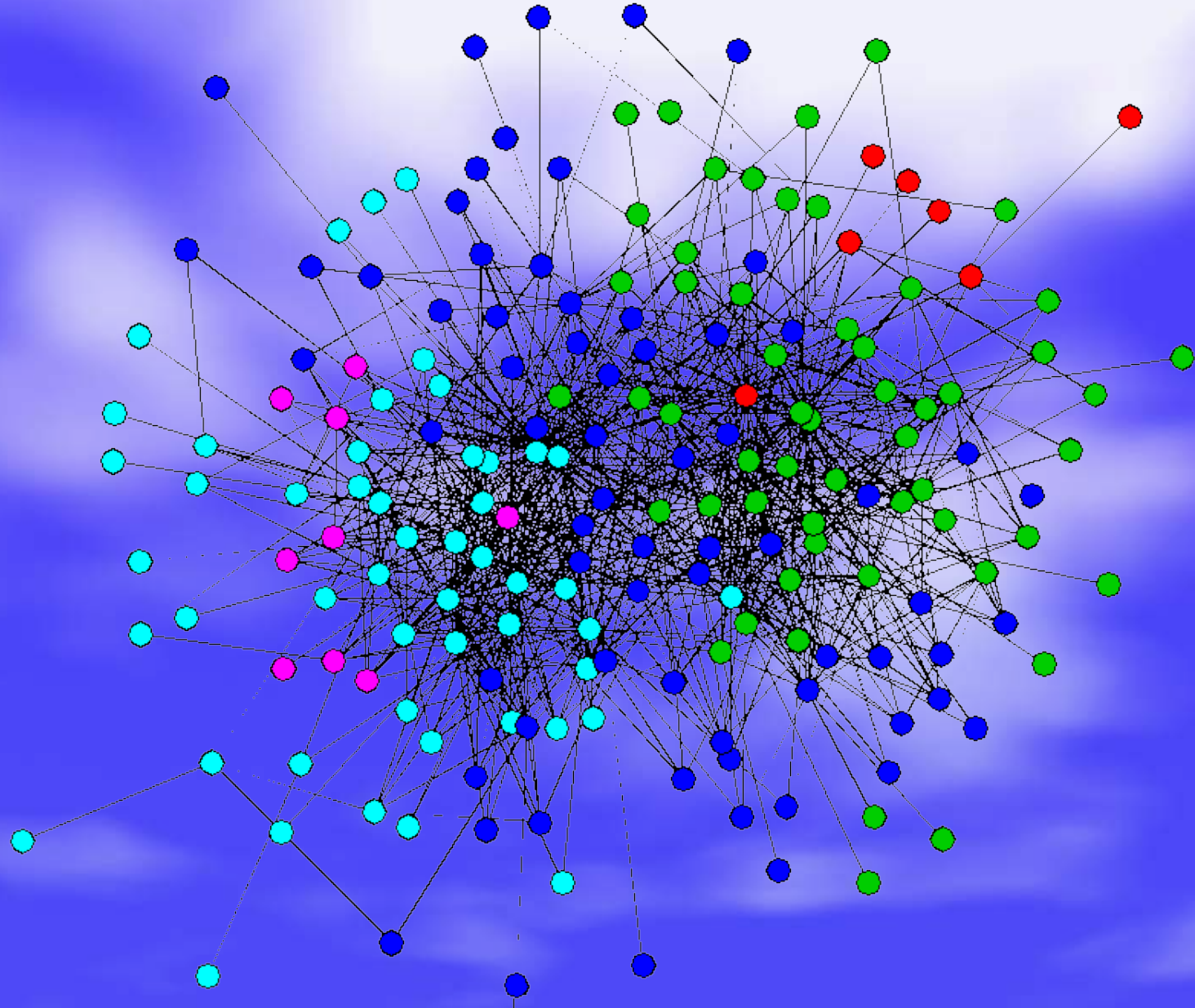
# Dataset

- Collected 250 URLs from yahoo directory's political blogs category
- Whereas others have sought specifically “liberal” and “conservative” sites, we included all
- Each site was hand annotated according to obviousness of bias, on a scale from 0 to 2 (Erin)

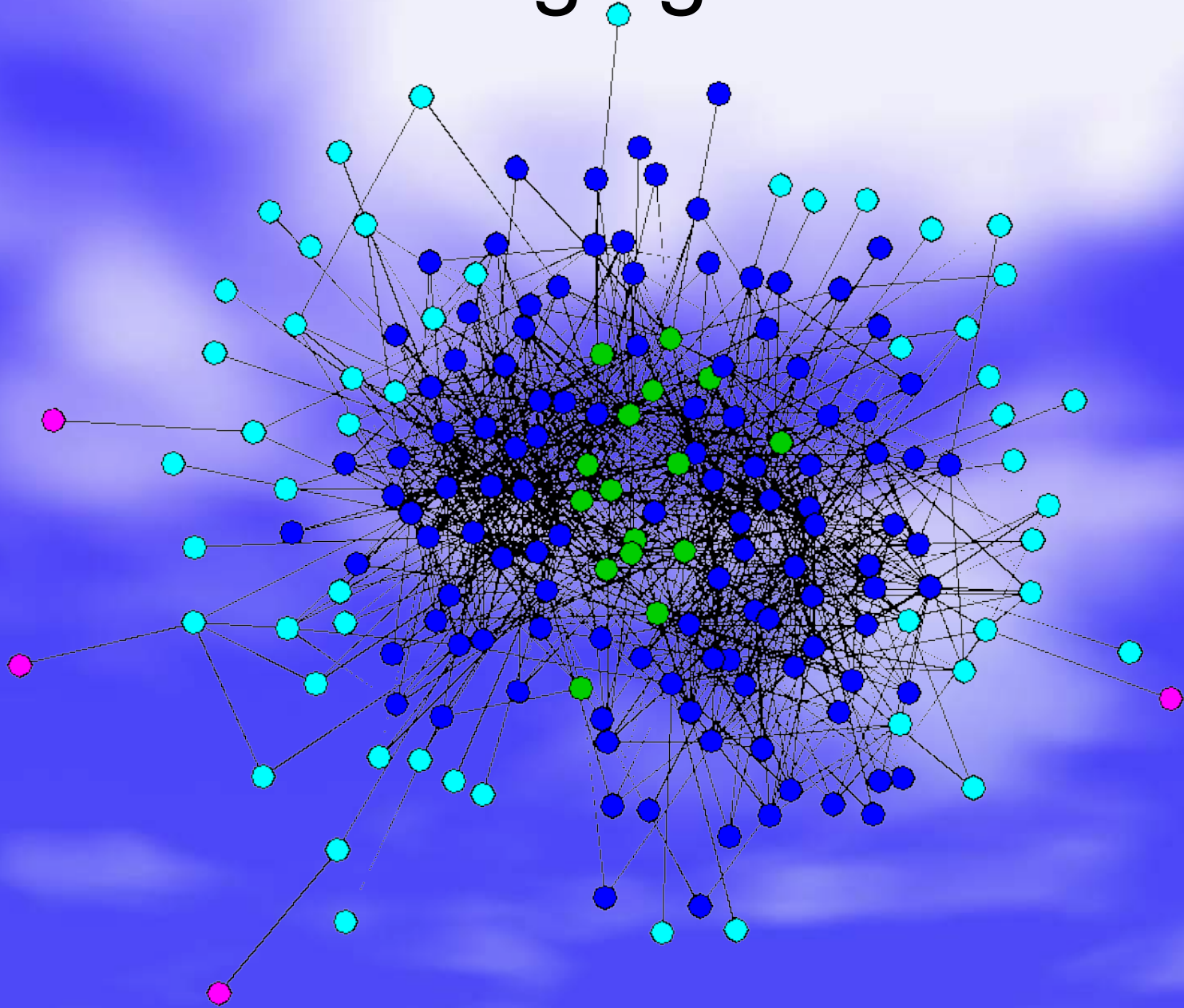
# Centrality



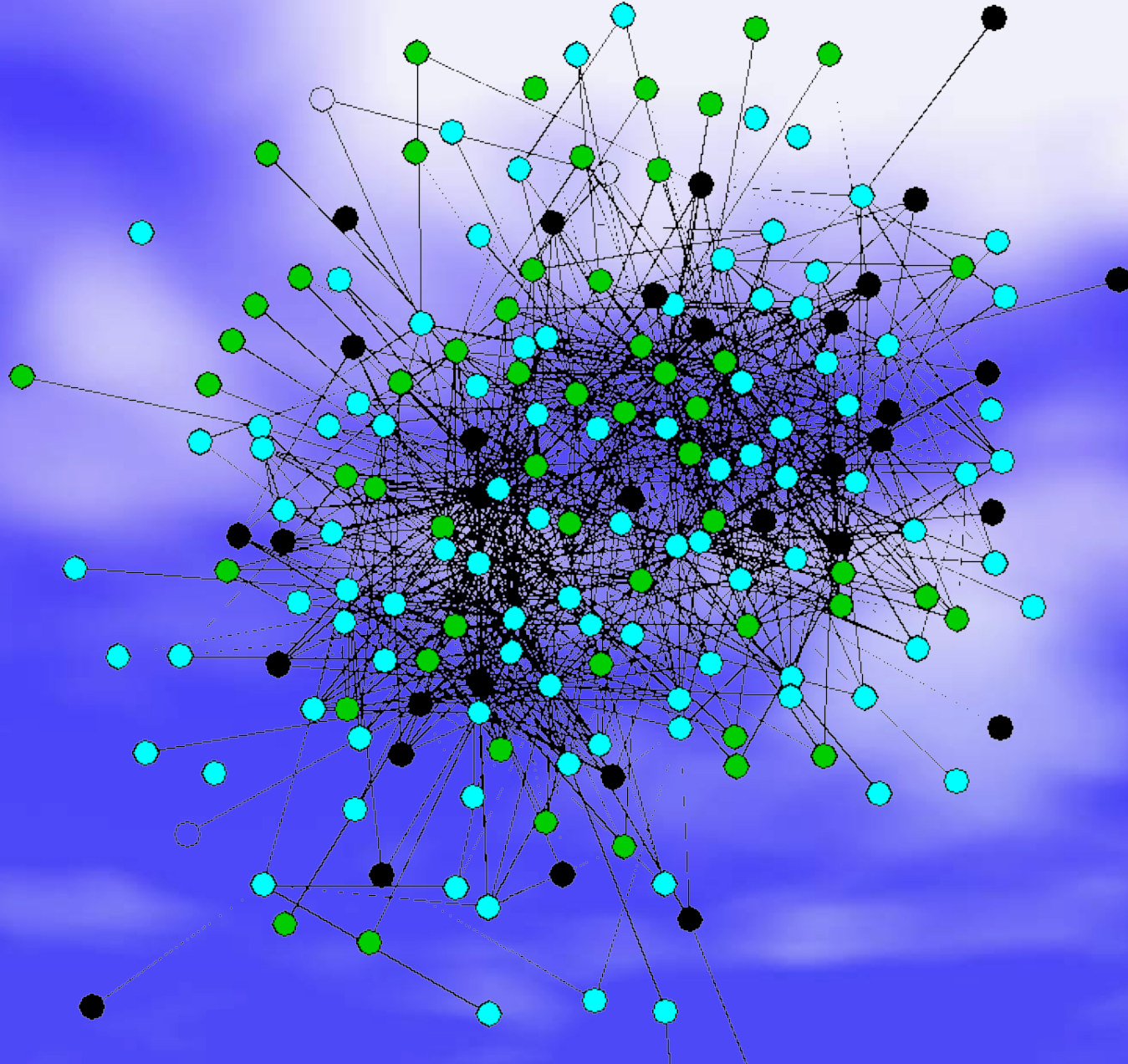
# Partisanship



# Bridging



# Perceived Bias



# Discussion

- No correlations found
- Partisanship and Bridging scales do still seem like useful information
- What could bias be?
  - Attention and selection bias: not aware or don't care about opposing data and views
  - Proportion and degree of stance taking: ratio of subjective to objective phrases; high/low polarity
  - Lots of covert stance taking (“weasel” words)

# Conclusion

- Need further sentiment analysis
  - Relationship between subjectivity and perceived bias?
  - Relationship between subjectivity and place in graph?
  - Distinction between overt and covert stance?
- Still missing first steps:
  - Identifying when a network is factionalized
  - Distinguishing topical structure from partisanship