

# Detecting Echo Chambers in social media; a graph-based approach

Francesco Zappia<sup>1</sup>  
Stefan Neumann<sup>1</sup> Aris Anagnostopoulos<sup>2</sup>  
Aris Gionis<sup>1</sup>

<sup>1</sup>KTH Royal Institute of Technology

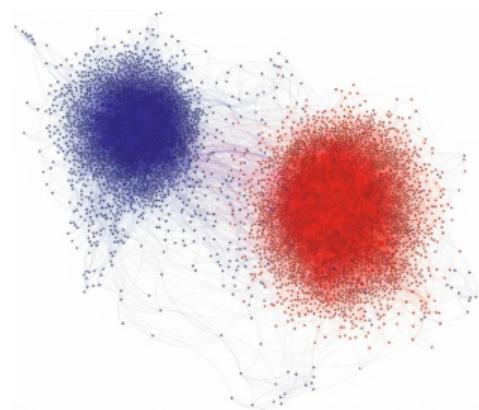
<sup>2</sup>Sapienza University of Rome

June 22, 2021



# Social media and Echo Chambers

- ▶ Social media are more and more diffused
- ▶ On these platforms we experience every day phenomena like polarization and **Echo Chambers**
- ▶ Echo chambers correspond to groups of users with the same ideas who, after interacting, reinforce their initial opinions



## Data model: Contents and Threads

Our definition of Echo Chamber is based on the concepts of **Contents** and **Threads**

- ▶ **Content:** the resource which is discussed. We can generally think of it as an online newspaper article. It is identified by a URL, e.g.

[https://www.breitbart.com/politics/2020/12/18/  
report-mark-zuckerbergs-.../](https://www.breitbart.com/politics/2020/12/18/report-mark-zuckerbergs-.../)

## Data model: Contents and Threads

Our definition of Echo Chamber is based on the concepts of **Contents** and **Threads**

- ▶ **Content:** the resource which is discussed. We can generally think of it as an online newspaper article. It is identified by a URL, e.g.

`https://www.breitbart.com/politics/2020/12/18/  
report-mark-zuckerbergs-.../`

- ▶ **Thread:** the social media *locality* discussing a content. Identified by its URL, e.g.

`https://twitter.com/DineshDSouza/status/1340726209446113280`

# Content and threads: an example

The image shows three side-by-side screenshots of a Twitter thread from the Breitbart News account (@BreitbartNews). The thread discusses a report alleging that Mark Zuckerberg and his wife gave \$419 million to nonprofits "improperly influenced" by the 2020 presidential election.

**Screenshot 1:** A tweet from Dinesh D'Souza (@DineshDSouza) at 8:21 AM on Sun, Dec 20, 2020. The tweet includes a photo of Mark Zuckerberg and the text: "Report Alleges Zuckerberg's \$419 Million 'Improperly Influenced Election'". It has 84 likes, 1,444 retweets, and 999 replies. A red circle highlights the like count (84).

**Screenshot 2:** A reply from Tom Fitton (@TomFitton\_1d) at 8:21 AM on Sun, Dec 20, 2020. The reply includes a photo of Mark Zuckerberg and the text: "Report Alleges Zuckerberg's \$419 Million 'Improperly Influenced Election'". It has 103 likes, 4,034 retweets, and 2,084 replies. A red circle highlights the like count (103).

**Screenshot 3:** A reply from Matt Lewis (@mattlewisinshow) at 8:21 AM on Sun, Dec 20, 2020. The reply includes a photo of Mark Zuckerberg and the text: "Report Alleges Zuckerberg's \$419 Million 'Improperly Influenced Election'". It has 1,061 likes, 11,264 retweets, and 4,387 replies. A red circle highlights the like count (1,061).

## Data model: Echo Chambers (1)

- ▶ We focus on **controversial contents**, i.e. contents that trigger a lot of hostile interactions *globally*. Common examples of controversial contents are politics and football

## Data model: Echo Chambers (1)

- ▶ We focus on **controversial contents**, i.e. contents that trigger a lot of hostile interactions *globally*. Common examples of controversial contents are politics and football
- ▶ We look for all the **threads associated to the content**, i.e. all discussions referencing the content URL

## Data model: Echo Chambers (1)

- ▶ We focus on **controversial contents**, i.e. contents that trigger a lot of hostile interactions *globally*. Common examples of controversial contents are politics and football
- ▶ We look for all the **threads associated to the content**, i.e. all discussions referencing the content URL

We can now give an initial informal definition of Echo Chambers.

- ▶ **Echo Chambers:** users discussing controversial contents with mostly friendly interactions

# Some graph concepts

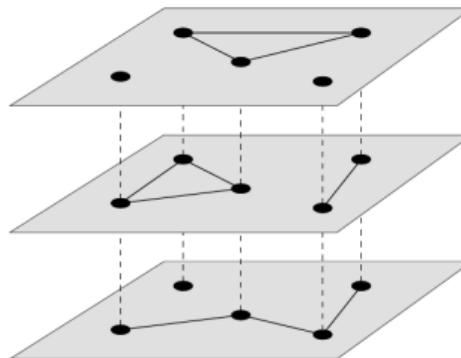
Before formalizing this idea, we need to introduce two concepts

- ▶ **Signed graphs:** graph with positive and negative edges

# Some graph concepts

Before formalizing this idea, we need to introduce two concepts

- ▶ **Signed graphs:** graph with positive and negative edges
- ▶ **Multiplex graph:** graph with multiple layers. Each layer has the same set of vertices but its own set of edges



## Data model: the interaction graph (1)

- ▶ We construct the **interaction graph**: a weighted signed multiplex graph
  - ▶ Vertices represent users

## Data model: the interaction graph (1)

- ▶ We construct the **interaction graph**: a weighted signed multiplex graph
  - ▶ Vertices represent users
  - ▶ Edges represent interactions between users

## Data model: the interaction graph (1)

- ▶ We construct the **interaction graph**: a weighted signed multiplex graph
  - ▶ Vertices represent users
  - ▶ Edges represent interactions between users
  - ▶ Edge weights  $w_e \in [-1, 1]$

## Data model: the interaction graph (1)

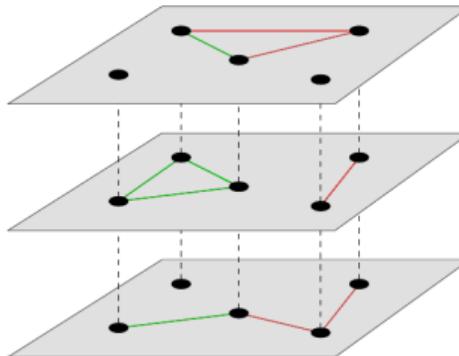
- ▶ We construct the **interaction graph**: a weighted signed multiplex graph
  - ▶ Vertices represent users
  - ▶ Edges represent interactions between users
  - ▶ Edge weights  $w_e \in [-1, 1]$
  - ▶ **Negative** edge: hostile interaction. **Positive** edge: friendly interaction

## Data model: the interaction graph (1)

- ▶ We construct the **interaction graph**: a weighted signed multiplex graph
  - ▶ Vertices represent users
  - ▶ Edges represent interactions between users
  - ▶ Edge weights  $w_e \in [-1, 1]$
  - ▶ **Negative** edge: hostile interaction. **Positive** edge: friendly interaction
  - ▶ Each layer is associated to a thread

# Data model: the interaction graph (1)

- We construct the **interaction graph**: a weighted signed multiplex graph
  - Vertices represent users
  - Edges represent interactions between users
  - Edge weights  $w_e \in [-1, 1]$
  - **Negative** edge: hostile interaction. **Positive** edge: friendly interaction
  - Each layer is associated to a thread



## Data model: Echo Chambers (2)

- ▶ **Echo Chambers:** users discussing controversial contents with mostly friendly interactions

## Data model: Echo Chambers (2)

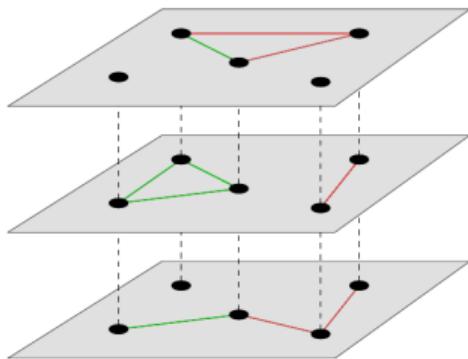
- ▶ **Echo Chambers:** users discussing controversial contents with mostly friendly interactions



- ▶ Consider only *non-controversial* threads associated to *controversial* contents
  - ▶ *Controversial:* more than  $\alpha$  percent of edges is negative

## Data model: $\mathcal{S}_C$

- ▶  $\mathcal{S}_C(U)$  contains threads which are *locally* non-controversial but it is defined only for contents that are *globally* controversial.



## Data model: Echo Chambers (3)

- ▶ **Echo Chambers:** users discussing controversial contents with mostly friendly interactions



- ▶ Consider only *non-controversial* threads associated to *controversial* contents

$$\xi(U) = \sum_{C \in \hat{\mathcal{C}}} \sum_{T[U] \in S_C(U)} (|T^+[U]| - |T^-[U]|)$$

- ▶  $T[U]$  is the subgraph induced by  $U$  in  $T$
- ▶  $|T^+[U]|$  is the number of positive edges in  $T[U]$

**Echo Chamber Problem:** find vertices  $U$  maximizing  $\xi(U)$ .

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem
- ▶ We show that it is hard to approximate this problem

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem
- ▶ We show that it is hard to approximate this problem
- ▶ We propose MIP and heuristic for solving it

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem
- ▶ We show that it is hard to approximate this problem
- ▶ We propose MIP and heuristic for solving it
- ▶ The heuristic achieves good performances on synthetic data

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem
- ▶ We show that it is hard to approximate this problem
- ▶ We propose MIP and heuristic for solving it
- ▶ The heuristic achieves good performances on synthetic data
- ▶ It has limitations on real-world data

# The Echo Chamber Problem: our results (1)

- ▶ We propose new methods for detecting echo chambers in social media: the Echo Chamber Problem
- ▶ We show that it is hard to approximate this problem
- ▶ We propose MIP and heuristic for solving it
- ▶ The heuristic achieves good performances on synthetic data
- ▶ It has limitations on real-world data

We will not show the proofs of our results.

## The Echo Chamber Problem: our results (2)

The Echo Chamber Problem cannot be approximated in polynomial time for a non-trivial factor.

### Theorem

*The ECP has no  $n^{1-\epsilon}$ -approximation algorithm for any  $\epsilon > 0$  unless  $\mathcal{P} = \mathcal{NP}$ .*

## The Echo Chamber Problem: our results (3)

An exact solution for the problem can be found through the following Mixed Integer Programming (MIP) Model

$$\text{maximize}_{T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}}} \sum_{ij \in E_k^+} \left( \sum x_{ij}^k - \sum_{ij \in E_k^-} x_{ij}^k \right) \quad (1)$$

$$\text{subject to} \quad x_{ij}^k \leq y_i \quad \forall ij \in E_k \quad (2)$$

$$x_{ij}^k \leq y_j \quad \forall ij \in E_k \quad (3)$$

$$x_{ij}^k \leq z_k \quad \forall ij \in E_k \quad (4)$$

$$x_{ij}^k \geq -2 + y_i + y_j + z_k \quad \forall ij \in E_k \quad (5)$$

$$\sum_{ij \in E_k^-} x_{ij}^k - \alpha \sum_{ij \in E_k} x_{ij}^k \leq 0 \quad \forall T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \quad (6)$$

$$y_i \in \{0, 1\} \quad \forall i \in V \quad (7)$$

$$0 \leq x_{ij}^k \leq 1 \quad \forall ij \in E_k \quad (8)$$

$$0 \leq z_k \leq 1 \quad \forall T_k \in \mathcal{T}_C, C \in \hat{\mathcal{C}} \quad (9)$$

# An heuristic for the Echo Chamber Problem

We propose a heuristic for solving the Echo Chamber Problem

- ▶ It is based on the relaxation of the MIP (integrality constraints are removed)
- ▶ Each edges is assigned a value in the MIP. This heuristic considers the vertices associated to edges with the highest value as possible solutions

# Echo Chamber Problem Validation

- ▶ No benchmark is available for such a problem

## Echo Chamber Problem Validation

- ▶ No benchmark is available for such a problem
- ▶ We use the Echo Chamber problem to find users in the same community

# Echo Chamber Problem Validation

- ▶ No benchmark is available for such a problem
- ▶ We use the Echo Chamber problem to find users in the same community



Clustering problem

## Data collection

- ▶ Graphs are built from Reddit or Twitter
- ▶ Similar data collection process:
  - ▶ Retrieve contents from posts of a  *subreddit* or of a Twitter account
  - ▶ Retrieve all threads related to that contents

# Data collection

- ▶ Graphs are built from Reddit or Twitter
- ▶ Similar data collection process:
  - ▶ Retrieve contents from posts of a  *subreddit* or of a Twitter account
  - ▶ Retrieve all threads related to that contents

The screenshot shows a vertical stack of three Reddit post cards. Each card has a blue upvote arrow icon, a numerical score, the poster's name, the post title, a snippet of the post, a thumbnail image, and a link preview.

- Post 1:** Posted by u/yaxxol 4 hours ago. Score: 3.9k. Title: Trump and his allies try to rewrite, distort history of pandemic while casting Fauci as public enemy No. 1. Preview: washingtonpost.com/politi...
- Post 2:** Posted by u/xRipleyx 3 hours ago. Score: 2.5k. Title: Progressive rep: 'Joe Manchin has become the new Mitch McConnell'. Preview: thehill.com/homene...
- Post 3:** Posted by u/Fr1sk3r 3 hours ago. Score: 2.8k. Title: Trump is a 'clear and present danger' and his words will 'surely kill again' says Lincoln Project co-founder. Preview: businessinsider.com/trump-...

Below the posts are standard Reddit navigation controls: back, forward, search, and refresh.

# Data collection

- ▶ Graphs are built from Reddit or Twitter
- ▶ Similar data collection process:
  - ▶ Retrieve contents from posts of a  *subreddit* or of a Twitter account
  - ▶ Retrieve all threads related to that contents
- ▶ Edge labels are obtained through a  *sentiment analyzer*

The image displays three separate screenshots of Reddit posts, each showing a different political topic with associated upvote/downvote counts, post times, and small thumbnail images.

- Post 1:** Posted by u/yaxxol 4 hours ago. Upvotes: 3.9k, Downvotes: 1. Content: "Trump and his allies try to rewrite, distort history of pandemic while casting Fauci as public enemy No. 1". Source: washingtonpost.com/politi... A thumbnail image shows two men in suits, one in a blue jacket and one in a dark suit.
- Post 2:** Posted by u/xRipleyx 3 hours ago. Upvotes: 2.5k. Content: "Progressive rep: 'Joe Manchin has become the new Mitch McConnell'". Source: thehill.com/homene... A thumbnail image shows two men, one in a blue shirt and one in a pink shirt, both looking towards the camera.
- Post 3:** Posted by u/Fr1sk3r 3 hours ago. Upvotes: 2.8k. Content: "Trump is a 'clear and present danger' and his words will 'surely kill again' says Lincoln Project co-founder". Source: businessinsider.com/trump-... A thumbnail image shows a portrait of Donald Trump.

## Data collection and generation

We obtained labeled datasets with different techniques

## Data collection and generation

We obtained labeled datasets with different techniques

- ▶ We generate synthetic data

# Data collection and generation

We obtained labeled datasets with different techniques

- ▶ We generate synthetic data
- ▶ r/asktrumpsupporters asks users to choose a flair among *Trump Supporter*, *Non Supporter* and *Undecided*

# Data collection and generation

We obtained labeled datasets with different techniques

- ▶ We generate synthetic data
- ▶ r/asktrumpsupporters asks users to choose a flair among *Trump Supporter*, *Non Supporter* and *Undecided*
- ▶ Twitter users are labeled according to the politicians they follow (if they are mostly *democrat* or *republican*). This dataset is built on @nytimes.

The image shows two Twitter profile cards side-by-side. The top card belongs to a user named 'DontCallMeMartha' who is a 'Nonsupporter'. The bottom card belongs to a user named 'Flussiges MOD' who is a 'Trump Supporter'. Both cards feature a small profile picture and a timestamp of '4m'.

User	Flair
DontCallMeMartha	Nonsupporter
Flussiges MOD	Trump Supporter

# Rounding algorithm evaluation

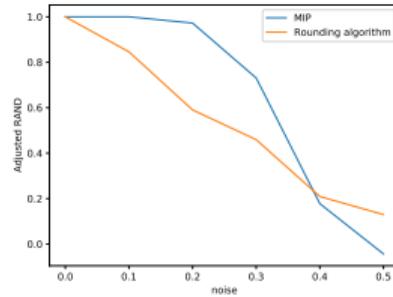
## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

# Rounding algorithm evaluation

## Observations on the results

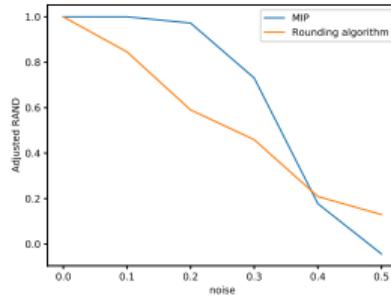
- ▶ The rounding algorithm is able to reconstruct communities in synthetic data



# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

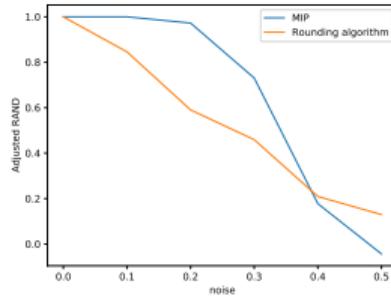


- ▶ It has limitations on real-world data (Adjusted RAND < 0.1)

# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

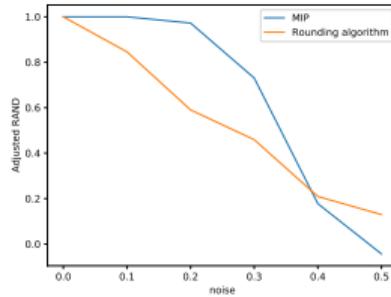


- ▶ It has limitations on real-world data ( $\text{Adjusted RAND} < 0.1$ )
- ▶ Different reasons

# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

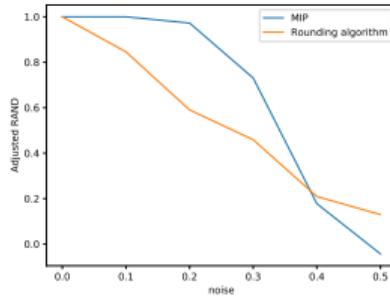


- ▶ It has limitations on real-world data ( $\text{Adjusted RAND} < 0.1$ )
- ▶ Different reasons
  - ▶ Non-validity of data model

# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

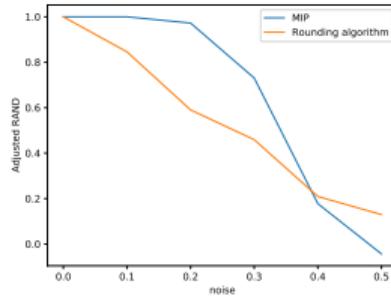


- ▶ It has limitations on real-world data ( $\text{Adjusted RAND} < 0.1$ )
- ▶ Different reasons
  - ▶ Non-validity of data model
  - ▶ Complexity of internet communication and sentiment analyzer

# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data

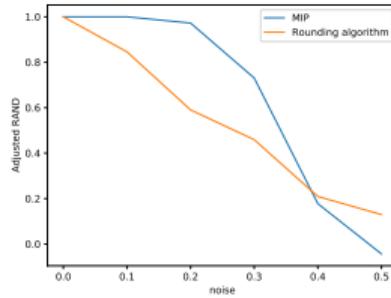


- ▶ It has limitations on real-world data ( $\text{Adjusted RAND} < 0.1$ )
- ▶ Different reasons
  - ▶ Non-validity of data model
  - ▶ Complexity of internet communication and sentiment analyzer
  - ▶ Limitations of the rounding algorithm

# Rounding algorithm evaluation

## Observations on the results

- ▶ The rounding algorithm is able to reconstruct communities in synthetic data



- ▶ It has limitations on real-world data ( $\text{Adjusted RAND} < 0.1$ )
- ▶ Different reasons
  - ▶ Non-validity of data model
  - ▶ Complexity of internet communication and sentiment analyzer
  - ▶ Limitations of the rounding algorithm
  - ▶ Sparsity of the data

# Conclusions

Our idea is that we may not be evaluating properly our heuristic

- ▶ The experiments on synthetic data show good performances
- ▶ Limitations are observed only on real-world data

# Conclusions

Our idea is that we may not be evaluating properly our heuristic

- ▶ The experiments on synthetic data show good performances
- ▶ Limitations are observed only on real-world data
- ▶ Need for a more accurate validation method

# Conclusions

Our idea is that we may not be evaluating properly our heuristic

- ▶ The experiments on synthetic data show good performances
- ▶ Limitations are observed only on real-world data
- ▶ Need for a more accurate validation method

Also, we propose new and alternative formulations

# Conclusions

Our idea is that we may not be evaluating properly our heuristic

- ▶ The experiments on synthetic data show good performances
- ▶ Limitations are observed only on real-world data
- ▶ Need for a more accurate validation method

Also, we propose new and alternative formulations

- ▶ Density can be taken into account

# Conclusions

Our idea is that we may not be evaluating properly our heuristic

- ▶ The experiments on synthetic data show good performances
- ▶ Limitations are observed only on real-world data
- ▶ Need for a more accurate validation method

Also, we propose new and alternative formulations

- ▶ Density can be taken into account
- ▶ Formulation can be extended with *follow graph*