Characterizing Information diffusion

Social Influence, Propagation Speed, Polarization

Giuseppe Manco

Institute of High Performance Computing and Networking

National Research Council

BEHAVIORAL MODELING
& SCALABLE ANALYTICS

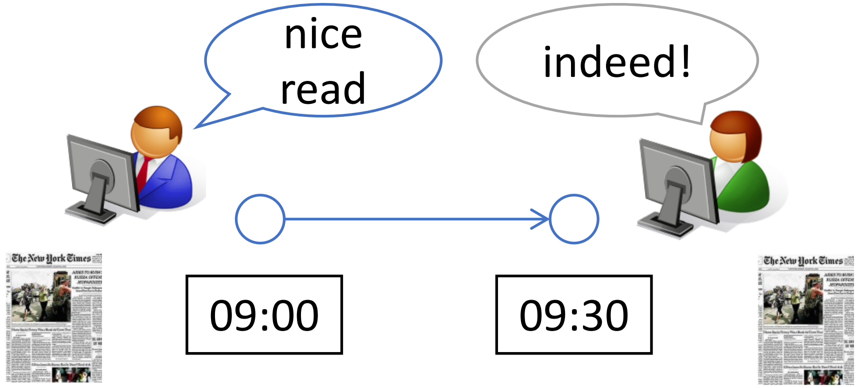
Disclaimer & Acknowledgements

- Summary of results of joint works with Francesco Bonchi, Nicola Barbieri, Corrado Monti, Ettore Ritacco

Agenda

- The Information Diffusion flow
- Influence and susceptibility
- Speed of propagation
- Polarity and echo chambers

Context
Information propagation in on-line social networks



users perform actions

post messages, pictures, video

buy, comment, link, rate, share, like, retweet

users are connected with other users

interact, influence each other

actions propagate

Relevant Questions

- What makes a content popular?
- Which creators are able to trigger a cascade?
- Who will share a content?
- When will someone share a content?
- Who is expert in a topic characterizing a set of contents?
- Who is interested in a topic?
- Which are the most popular topics?
- …

Basic Notation

- A graph G = (V,E)
  - Users u,v in V
- Items i
  - A cascade all (time-sorted) users who adopt that item

Node	Action	Time
u_1	i_1	t_0
u_2	i_1	t_1
u_3	i_2	t_3
…	…	…

- $y_{i,u}$ binary indicator

  u adopts i

A Basic Tool: Probabilistic modeling

- Treat data as observations that arise from a generative probabilistic process that includes hidden variables
  - For cascade data, the hidden variables reflect the behavior of single users or the commonality in actions
- Infer the hidden structure using posterior inference
  - What are the latent behaviors that describe this group of users?
- Situate new data into the estimated model.
  - How does a new user fit into the estimated behavioral structure?

Key Insight 1
Propagations are topic-dependent

Barbieri, Bonchi, Manco: Topic Aware social influence propagation models, ICDM2012

Key insight 1:
Propagations are topic-aware

Generalization of the IC Model: the probability of a user become active depends on the strength exerted by one of its precedessors, which in turn depends on the **topic k of interest**.

$$\mathcal{L}(i,u) = y_{i,y} \cdot \log\left(\sum_k P_{u,k}^{i,+}\right) + (1 - y_{i,u}) \cdot \sum_k \log P_{u,k}^{i,-}$$

Probability that some of the potential influencers in activating u

Probability that none of the the “out-of-react” influencers succeeds in activating u

$$P_{u,k}^{i,+} = 1 - \prod_v (1 - p_{u,v}^k)$$

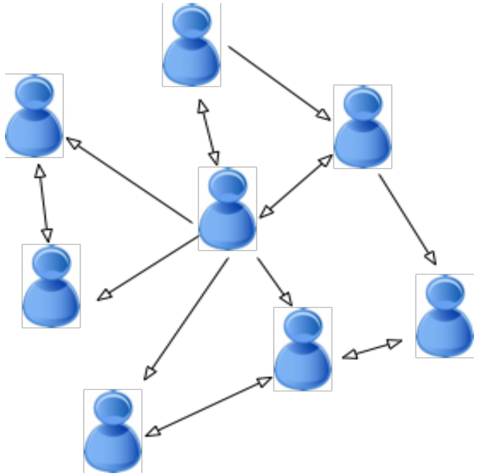
$$P_{u,k}^{i,-} = \prod_v (1 - p_{u,v}^k)$$

Nicola Barbieri, Francesco Bonchi, Giuseppe Manco: Topic-aware social influence propagation models. Knowl. Inf. Syst. 37(3): 555-584 (2013)

Topic-aware Social Influence Propagation Models

Input:

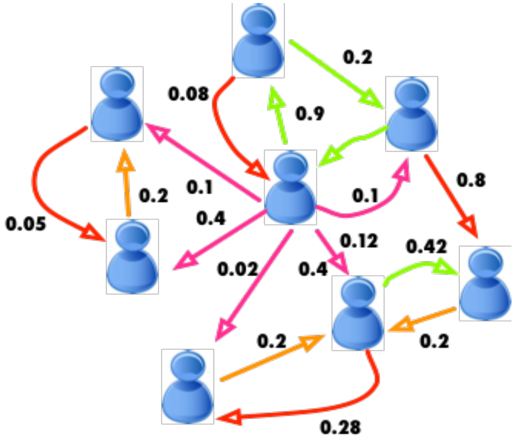
Social Network + Propagation Log
+ number of topics K



Node	Action	Time
u_1	i_1	t_0
u_2	i_1	t_1
u_3	i_2	t_3
…	…	…

Output:

Topics +
Topic-Aware propagation
Strengths



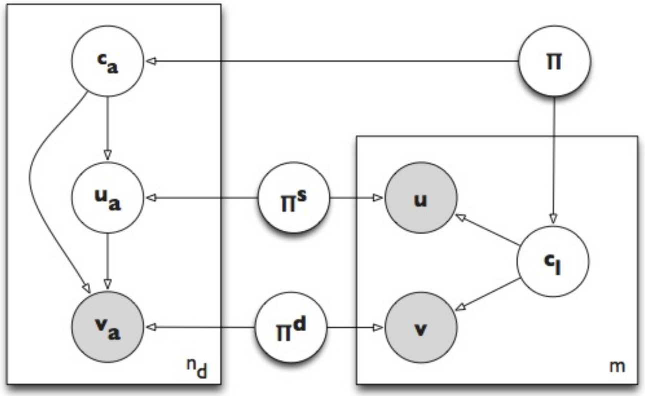
Better users' behavior modeling

Applications:

Detection of influent authorities for different topics

Design of targeted viral marketing strategies

Key insight 2
Adoptions and connections depend on influence



Probability of a link

$$P((u,v) \in E) = \vartheta_{u,k} \cdot \varphi_{v,k}$$

(source) $\vartheta_u = SoftMax(\Pi^s)$

(destination) $\varphi_v = SoftMax(\Pi^d)$

Probability of an action being propagated

$$p^k_{u,v} \propto \theta_{u,k} \cdot \phi_{k,v}$$

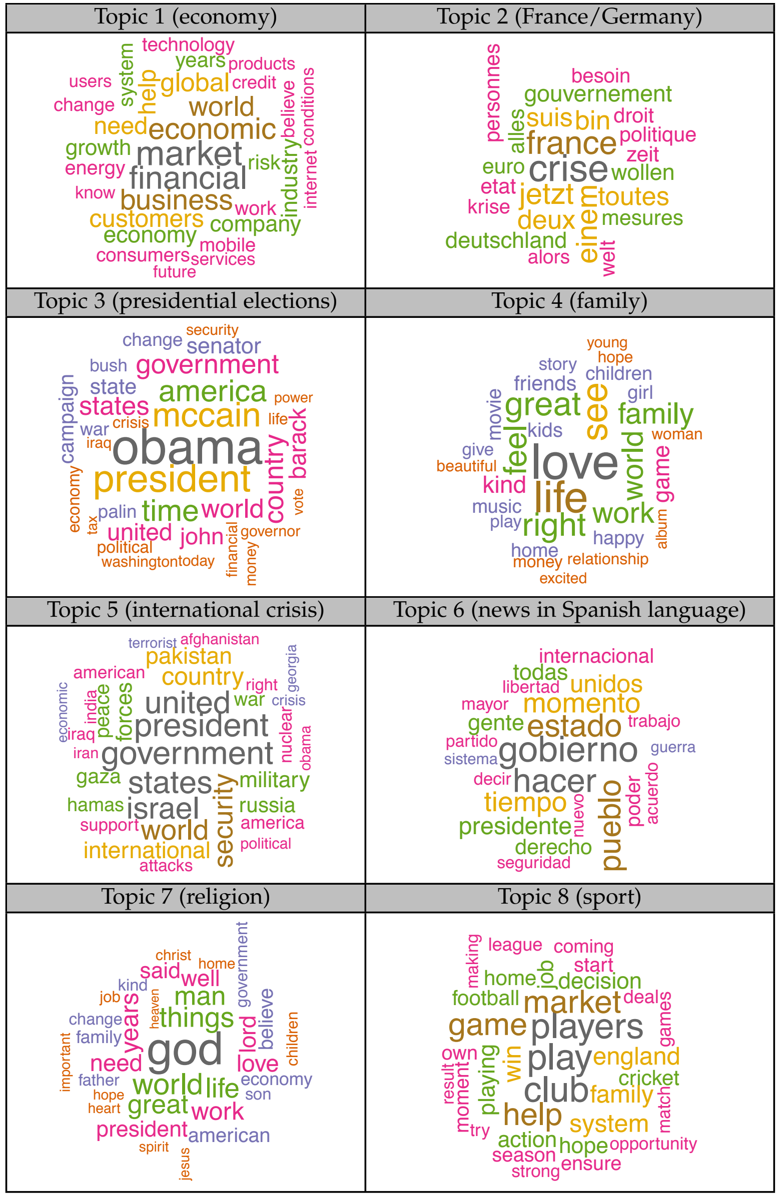
(influencer) $\varphi_v = SoftMax(\Pi^s)$

(susceptible) $\theta_u = SoftMax(\Pi^d)$

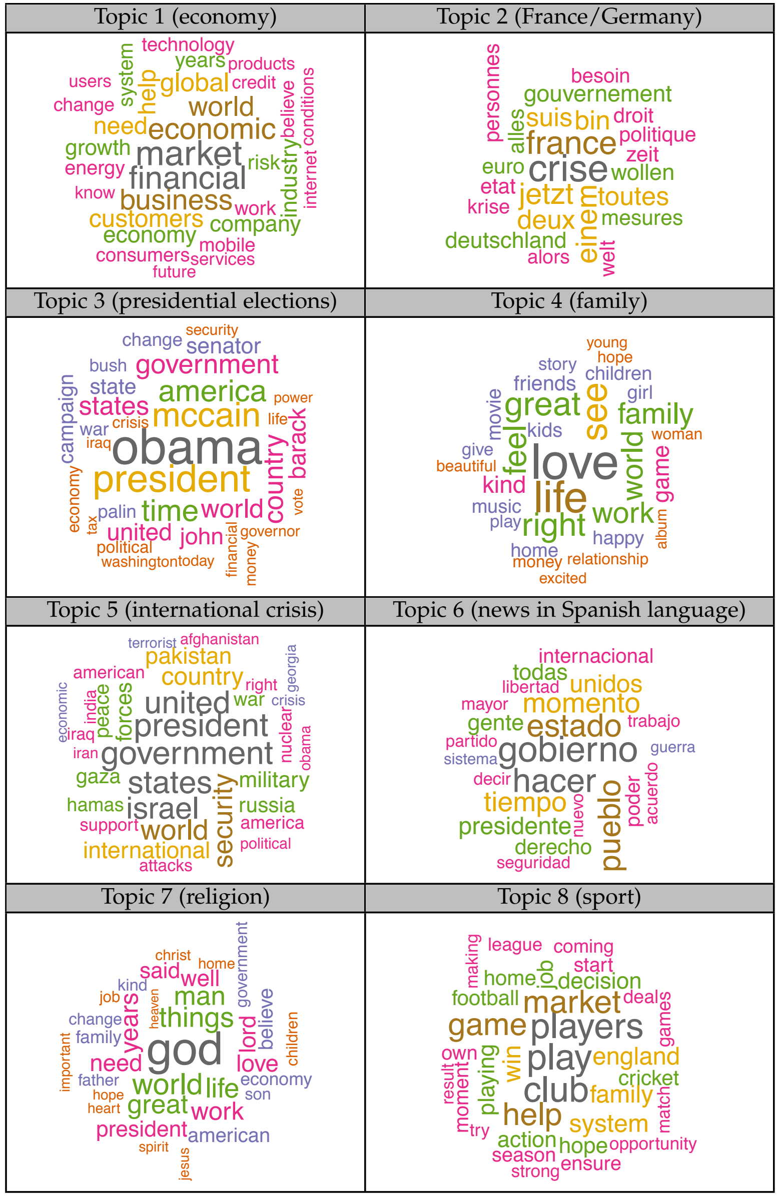
Nicola Barbieri, Francesco Bonchi, Giuseppe Manco: Cascade-based community detection. WSDM 2013: 33-42

Nicola Barbieri, Francesco Bonchi, Giuseppe Manco: Who to follow and why: link prediction with explanations. KDD 2014: 1266-1275

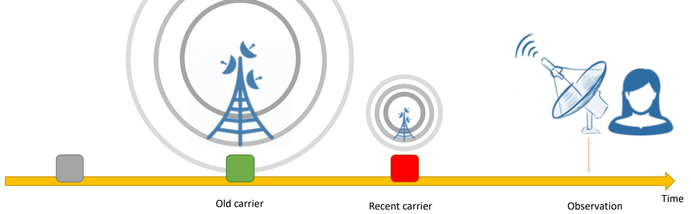
Topics on Memetracker



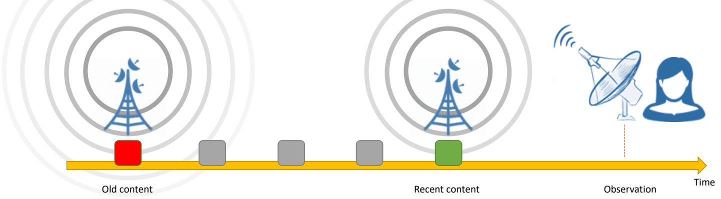
Influential nodes on Memetracker



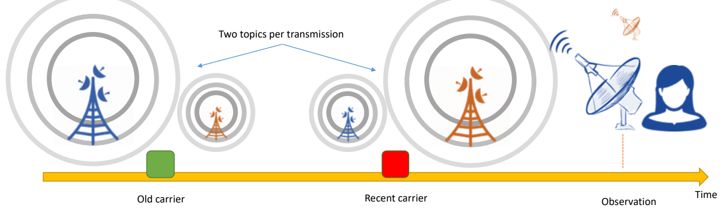
- **Contagion is carrier-dependent**
  - Some carriers are more infectious than others
    - **Influence** exerted in the diffusion process



- **Contagion is time-dependent**
  - the probability of contagion depends on the time when the target gets in touch with the content



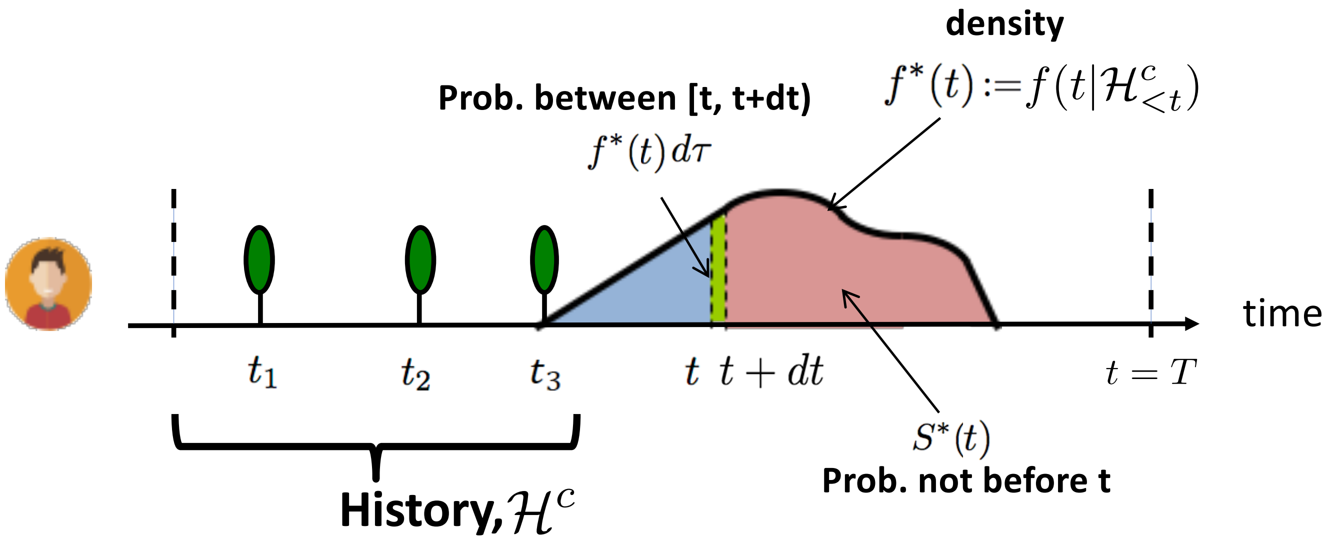
- **Contagion is topic-dependent**
  - **Susceptibility** and **influence** are relative the content
  - Content is characterized by **topics**



Modeling the contagion process

Key insight 3
The speed of propagations depend on influence

Modeling time as a random variable
Survival Networks



Intensity:

Probability between [t, t+dt) but not before t

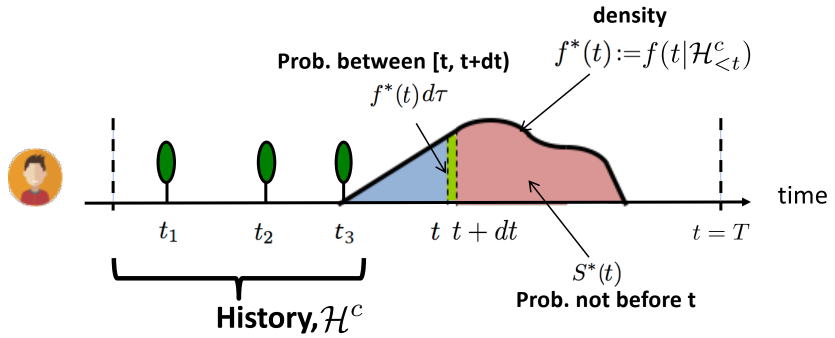
$$\lambda^*(t)dt = \frac{f^*(t)dt}{S^*(t)} \geq 0$$

$\lambda^*(t)$ **It is a rate = # of events / unit of time**

Source: Gomez-Rodriguez, Valera: Learning with Temporal Point Processes, ICML2018

Modeling time as a random variable
Survival Networks

Idea: model the intensity function as the effect of influence and susceptibility



Intensity:
Probability between [t, t+dt) but not before t

$$\lambda^*(t)dt = \frac{f^*(t)dt}{S^*(t)} \geq 0$$

$\lambda^*(t)$ **It is a rate = # of events / unit of time**

$$\lambda(t_i|u, v, k) \propto \theta_{u,k} \cdot \phi_{u,k}$$

Node	Action	Time
u_1	i_1	t_0
u_2	i_1	t_1
u_3	i_2	t_3
…	…	…

Nicola Barbieri, Giuseppe Manco, Ettore Ritacco: Survival Factorization on Diffusion Networks. TKDE 2019
Giuseppe Manco, Giuseppe Pirrò, Ettore Ritacco: Predicting Temporal Activation Patterns via Recurrent Neural Networks. ISMIS 2018: 347-356

Summary so far

It’s all about representation learning!

Authoritativeness of a user in a topic: $\varphi_{u,k}$

Interest of a user for a topic: $\theta_{u,k}$

Relevance of an item for a topic: $\gamma_{i,k}$

Lady Gaga @ladygaga

Justin Bieber @justinbieber

Barack Obama @barackobama

CNN @cnn

The Economist @TheEconomist

0.92 0.08

0.01 0.09

Key insight 4 Polarization and ideological leanings

- We can also use latent representations for modeling polarization
  - Authoritativeness of a user in a topic: $\varphi_{u,k}$
  - Interest of a user for a topic: $\theta_{u,k}$
  - Relevance of an item for a topic: $\gamma_{i,k}$
- Polarization

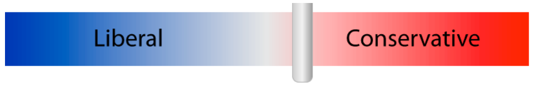
$$\ell_{u,k} \in \{+, -\}$$

- Probability of positive polarization

$$\phi_{u,k} \triangleq P(\ell_{u,k} = +)$$

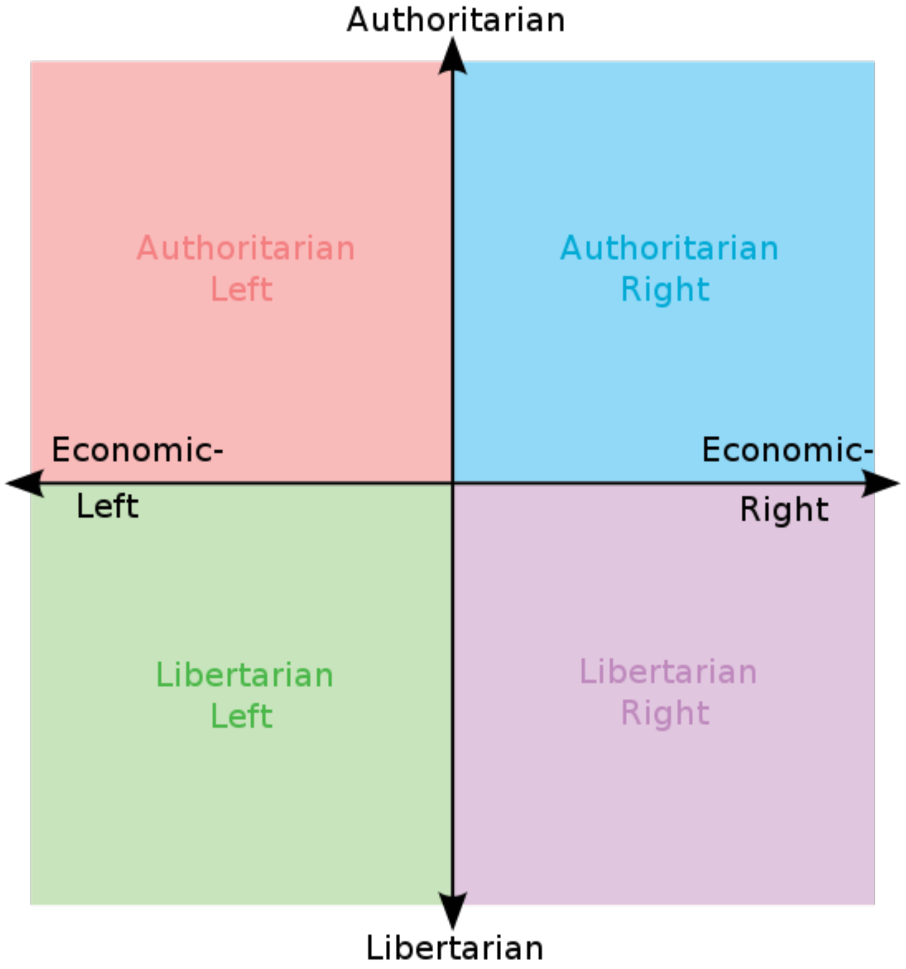
Polarization and ideological leanings

- Many works in opinion modeling assume a single axis
  - Often ill-defined too



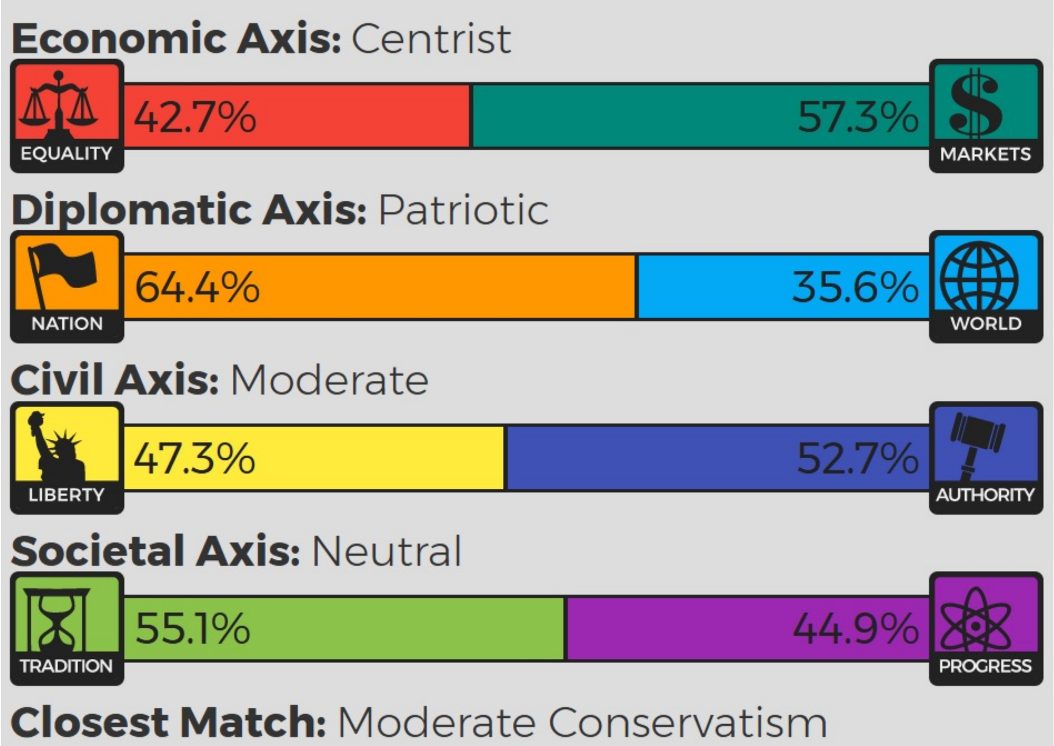
Why ideological embeddings?

- Multi-dimensional ideological spaces



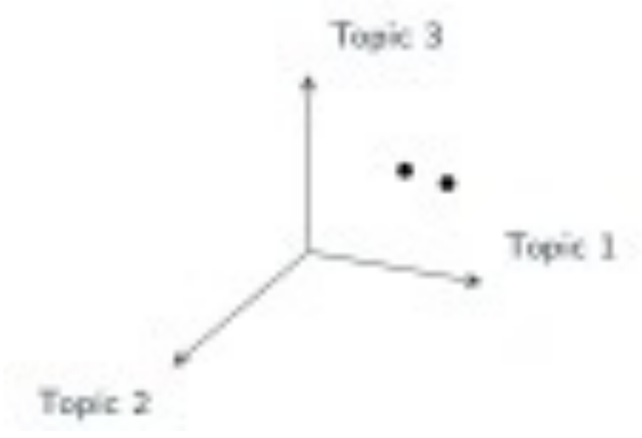
Why ideological embeddings?

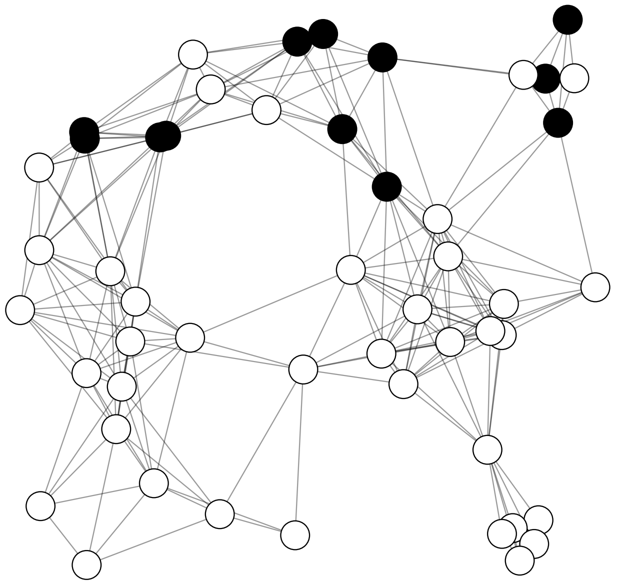
- Multi-dimensional ideological spaces



Which behavior makes this observable?

- Spreading news on social media
  - A graph of agents that spread news content
  - They spread the same item if they are ideologically aligned





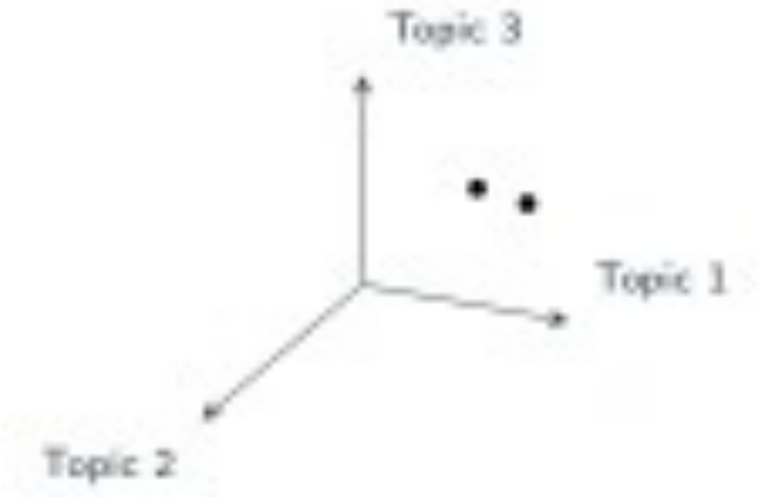
Which behavior makes this observable?

- We model actions that produce cascades on a graph through homophily:

a node imitates a neighbor
if they are similar in ideological
space.

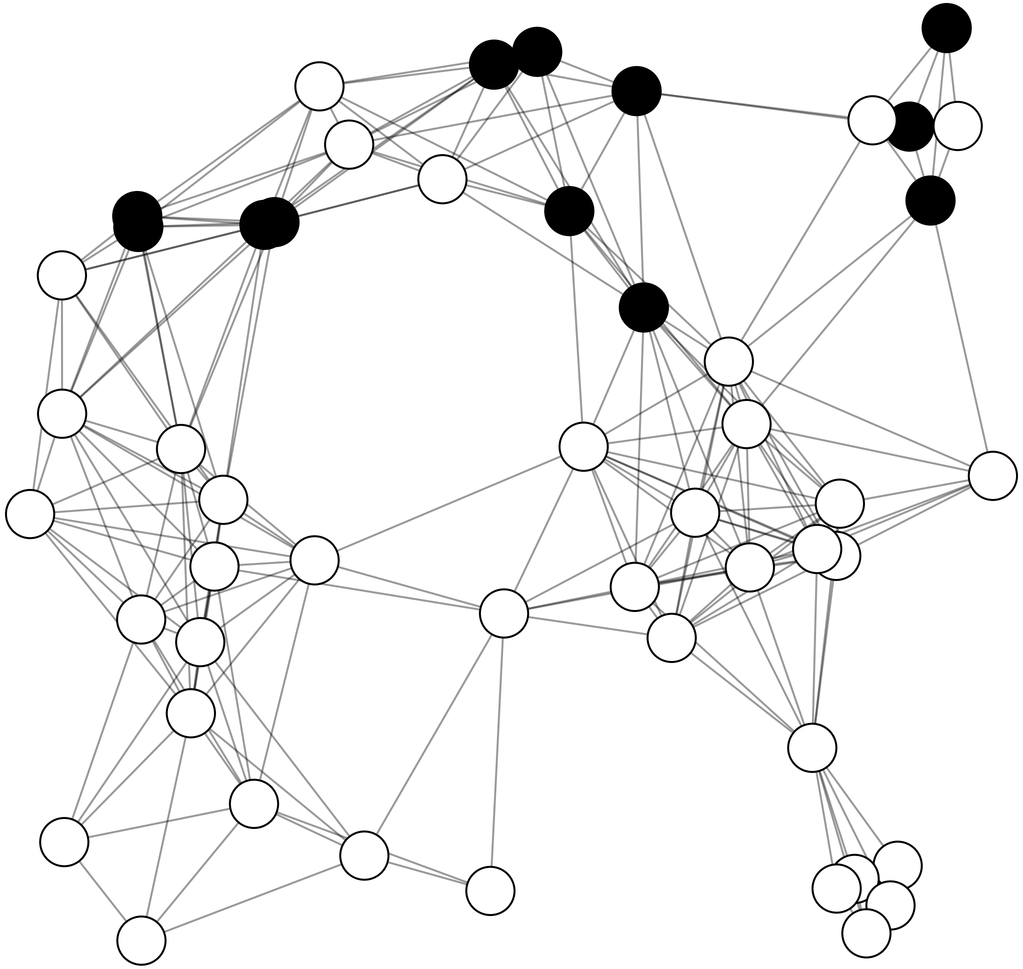
What is an ideological space?

- User-defined
- Each item must be tagged
  - a news can be “healthcare’ or “migration” • Manually or LDA
  - Can be fuzzy



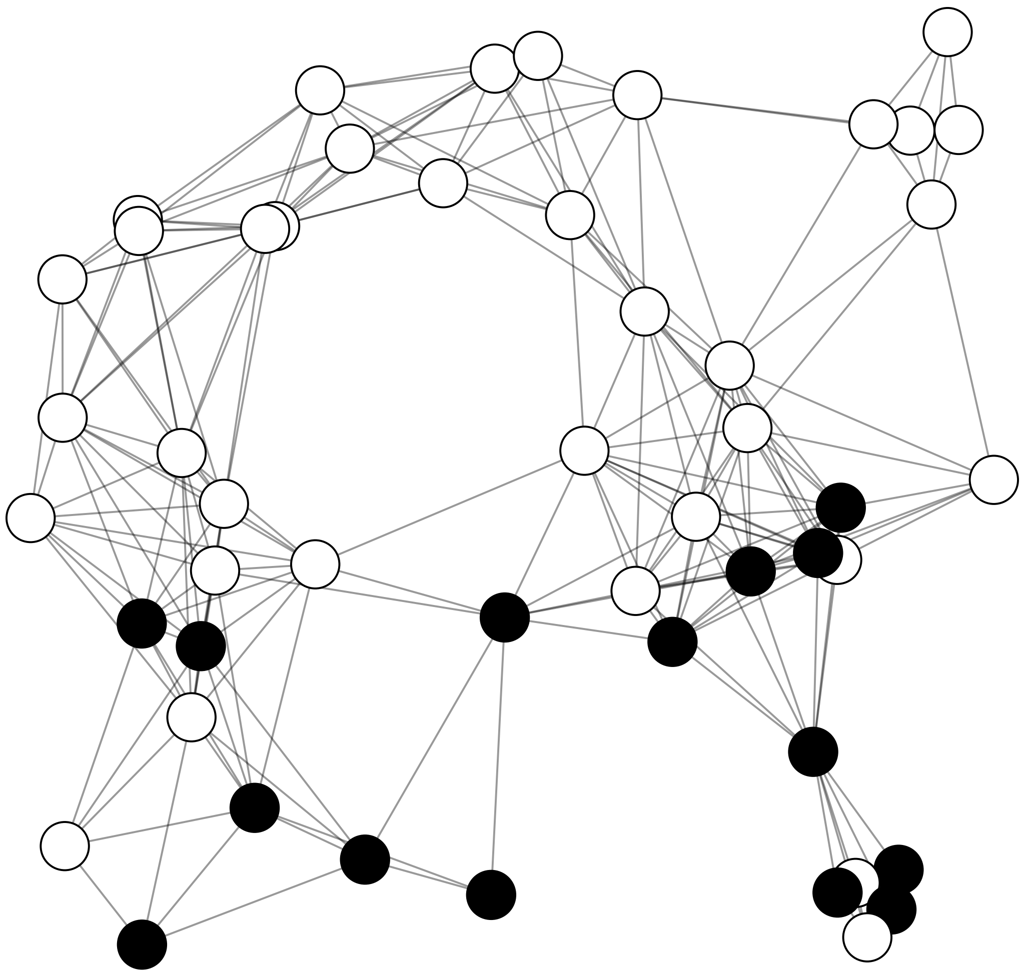
○ = user, ● = user sharing an item

Item 1
0.9 healthcare + 0.1 migration



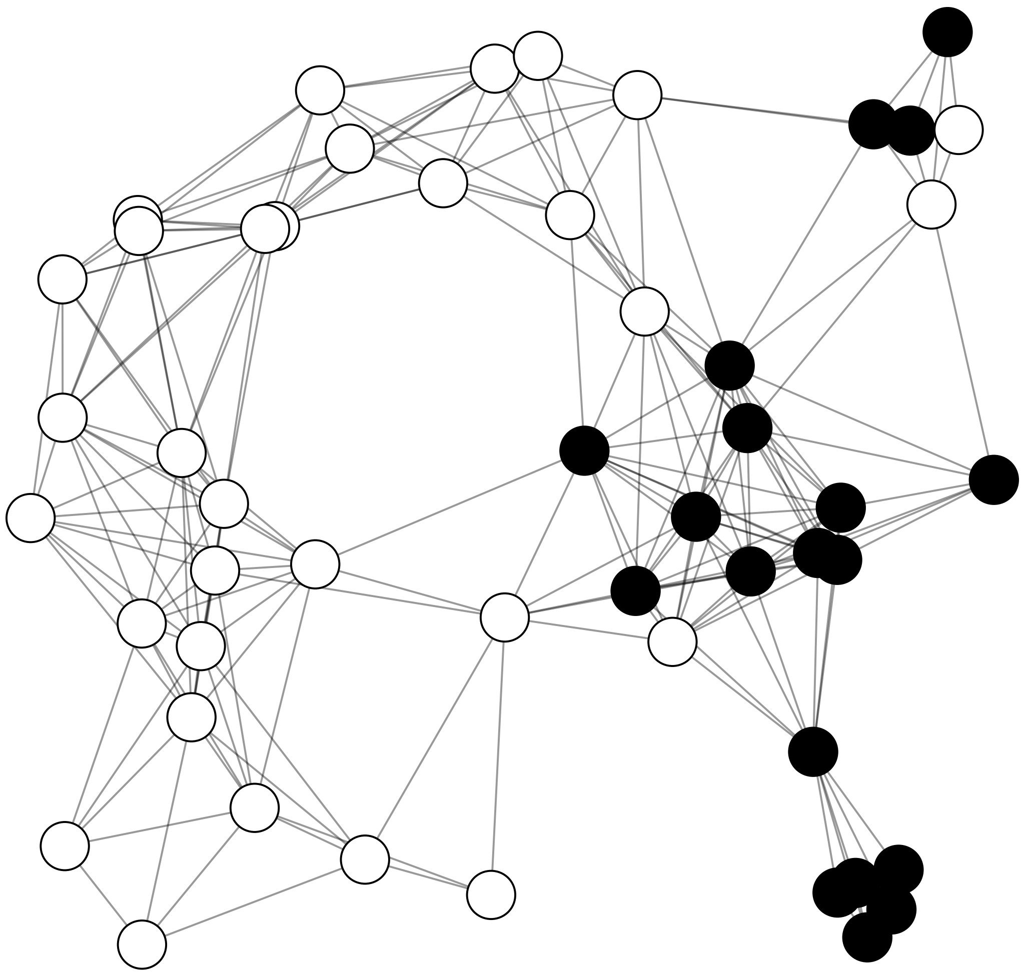
○ = user, ● = user sharing an item

Item 2
0.9 healthcare + 0.1 migration



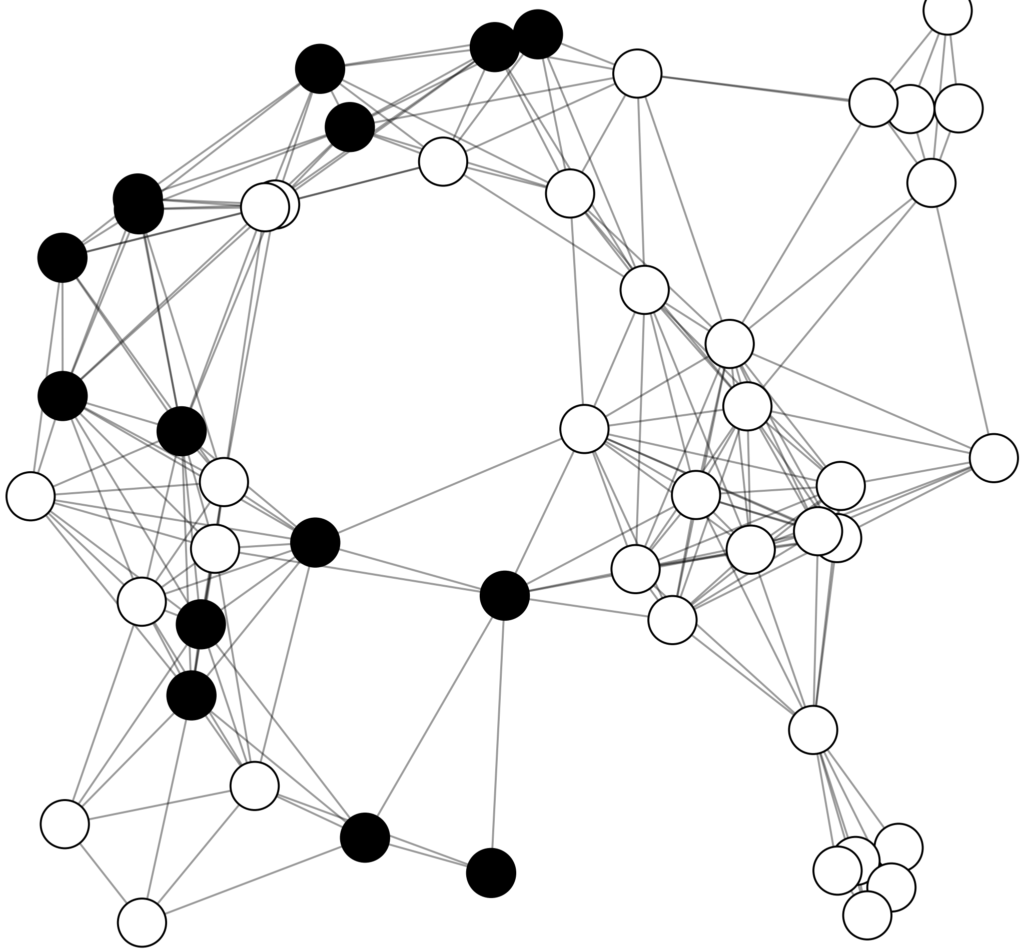
○ = user, ● = user sharing an item

Item 3
0.1 healthcare + 0.9 migration

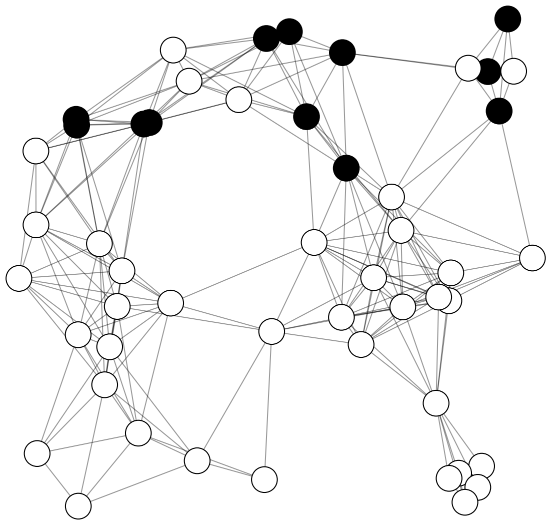


○ = user, ● = user sharing an item

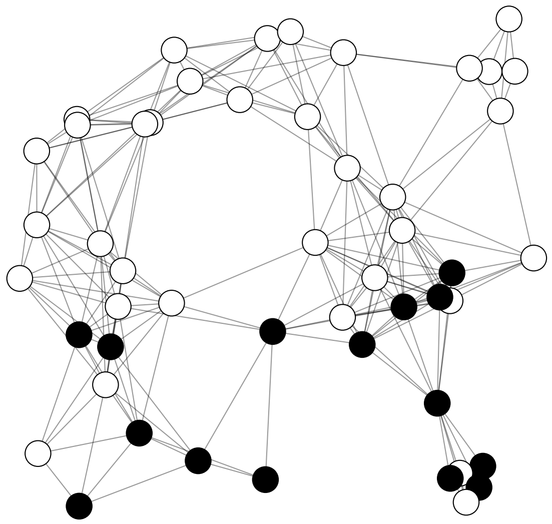
Item 4
0.1 healthcare + 0.9 migration



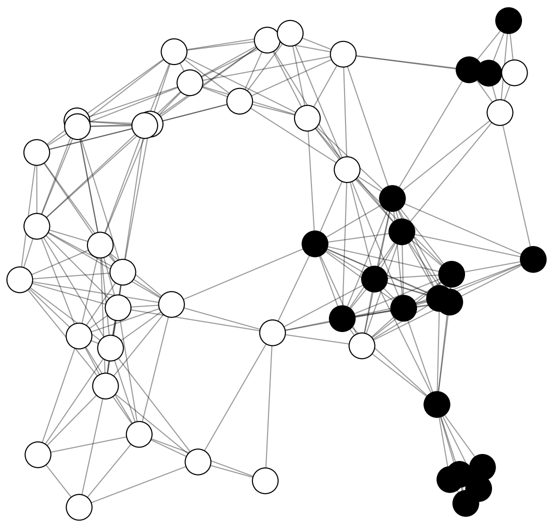
Item 1
0.9 healthcare + 0.1 migration



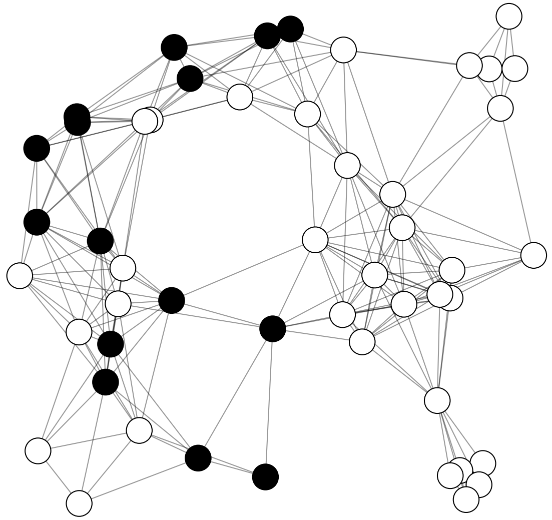
Item 2
0.9 healthcare + 0.1 migration

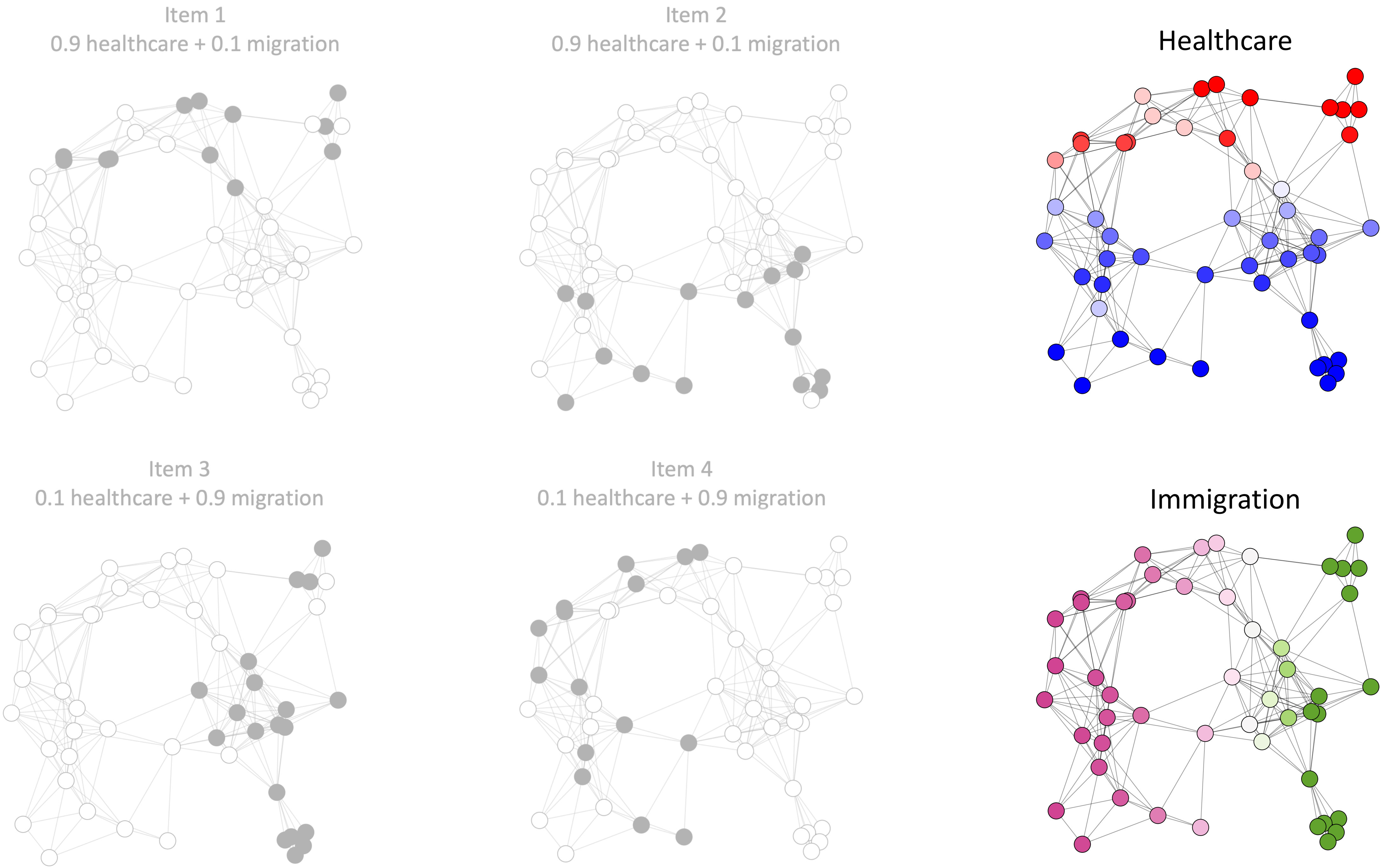


Item 3
0.1 healthcare + 0.9 migration



Item 4
0.1 healthcare + 0.9 migration




Item 1
0.9 healthcare + 0.1 migration
Item 2
0.9 healthcare + 0.1 migration
Healthcare
Item 3
0.1 healthcare + 0.9 migration
Item 4
0.1 healthcare + 0.9 migration
Immigration

The Generative model

- Interests $\vartheta_{u,k} \sim \text{Beta}(\alpha, \beta)$
- Polarities $\phi_{u,k} \sim \text{Beta}(p^{-1}, p^{-1})$
- Generate an item i:
  - topic distribution $\gamma_i \sim \text{Dir}(\mathbf{q})$
  - initial activator $v_i \sim U(V)$
  - For $(v, u) \in E$ s.t. v activated and u has not seen i:
    1. Node u sees i from v
    2. topic $k \sim \text{Multinom}(\gamma_i)$
    3. u is interested with probability $\vartheta_{u,k}$
    4. If interested, attitudes of u and v are Bernoullis with $\phi_{u,k}$ and $\phi_{v,k}$
    5. If equal attitudes, u activates on i

Modeling diffusion as ideological alignment

$$\mathcal{L}(i,u) \approx \sum_{v} \left\{ y_{i,u} \cdot \log\left(P_{u,v}^{i,+}\right) + (1 - u_{i,u}) \cdot \log\left(1 - P_{u}^{i,+}\right) \right\}$$

$$P_{u,v}^{i,+} = \sum_{k} \gamma_{i,k} \cdot \theta_{u,k} \cdot p(u,v,k)$$

Relevance

Interest

Alignment in polarities

Polarity alignment

$$p(u, v, k) = \underbrace{\phi_{u,k} \cdot \phi_{v,k}}_{\text{Positive alignment}} + \underbrace{(1 - \phi_{u,k}) \cdot (1 - \phi_{v,k})}_{\text{Negative alignment}}$$

Learning the model

$$\mathcal{L}(i,u) \approx \sum_{v} \left\{ y_{i,u} \cdot \log\left(P^{i,+}_{u,v}\right) + (1 - u_{i,u}) \cdot \log\left(1 - P^{i,+}_{u}\right) \right\}$$

- Fast and effective gradient-based optimization
  - Each potential activation u → v is an example
  - Scales to a large number of cascades
- Extensible to support variational inference

Algorithm 1 Inference algorithm.

Input:

- Graph (V, E)
- For every item $i \in \mathcal{I}$:
  - Its topic distribution γ_i on K topics
  - Its activated nodes $\mathcal{D}_i \subseteq V$

Output: Polarities ϕ_u and interests θ_u for all $u \in V$.

Initialize ϕ and θ as $|V| \times K$ matrices.
for number of epochs **do**
    for $i \in \mathcal{I}$ **do**
        for $v \in \mathcal{D}_i$ **do**
            for $u \in \{u \in \mathcal{D}_i | (v, u) \in E\}$ **do**
                Update ϕ, θ by ascending the gradient:

$$\nabla_{\phi,\theta} \log\left(\sum_k \gamma_{i,k} \cdot \theta_{u,k} \cdot p(u, v, k)\right)$$

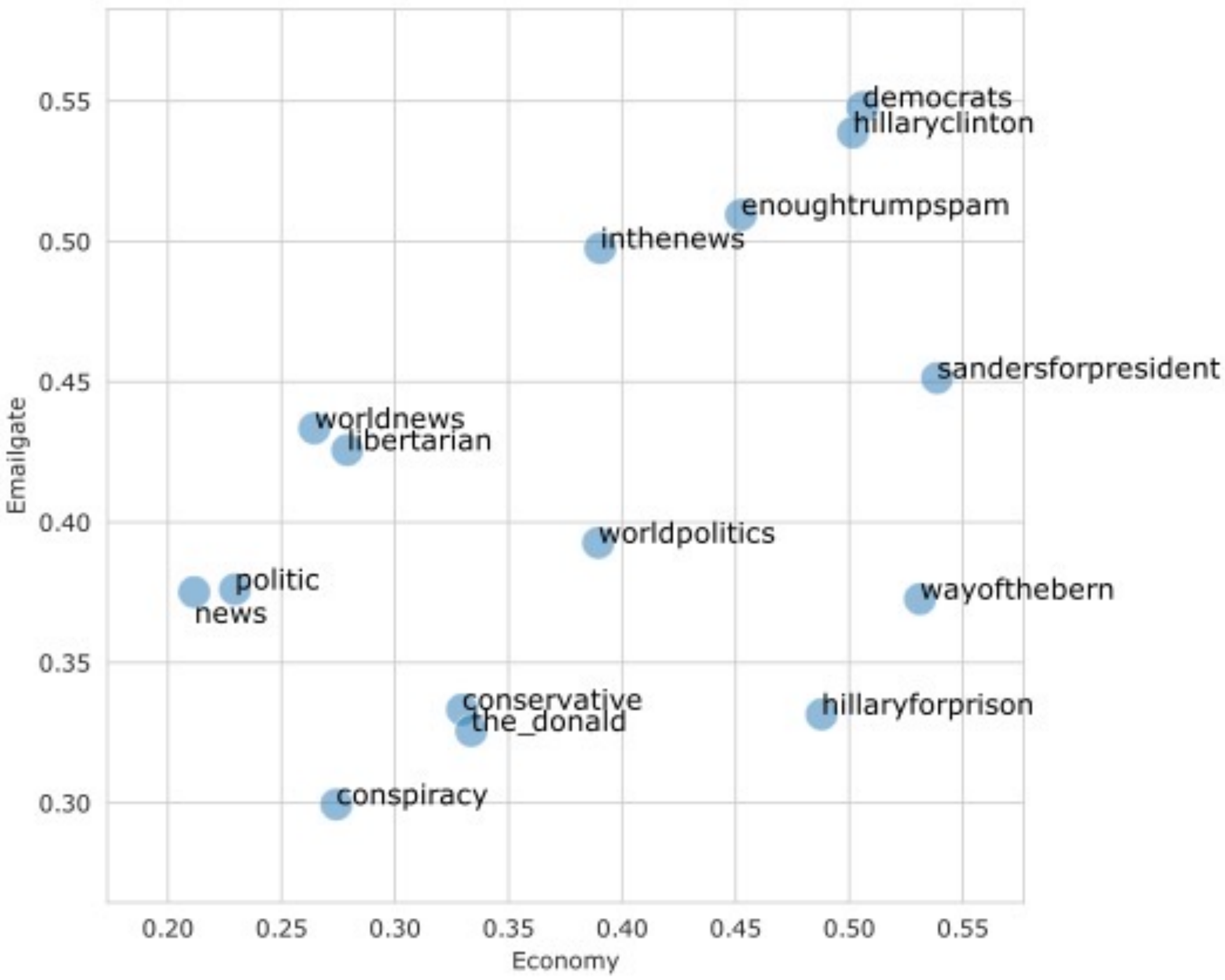
            end for
            for $u \in \textsc{sample}(\{u \notin \mathcal{D}_i | (v, u) \in E\})$ **do**
                Update ϕ, θ by ascending the gradient:

$$\nabla_{\phi,\theta} \log\left(1 - \sum_k \gamma_{i,k} \cdot \theta_{u,k} \cdot p(u, v, k)\right)$$

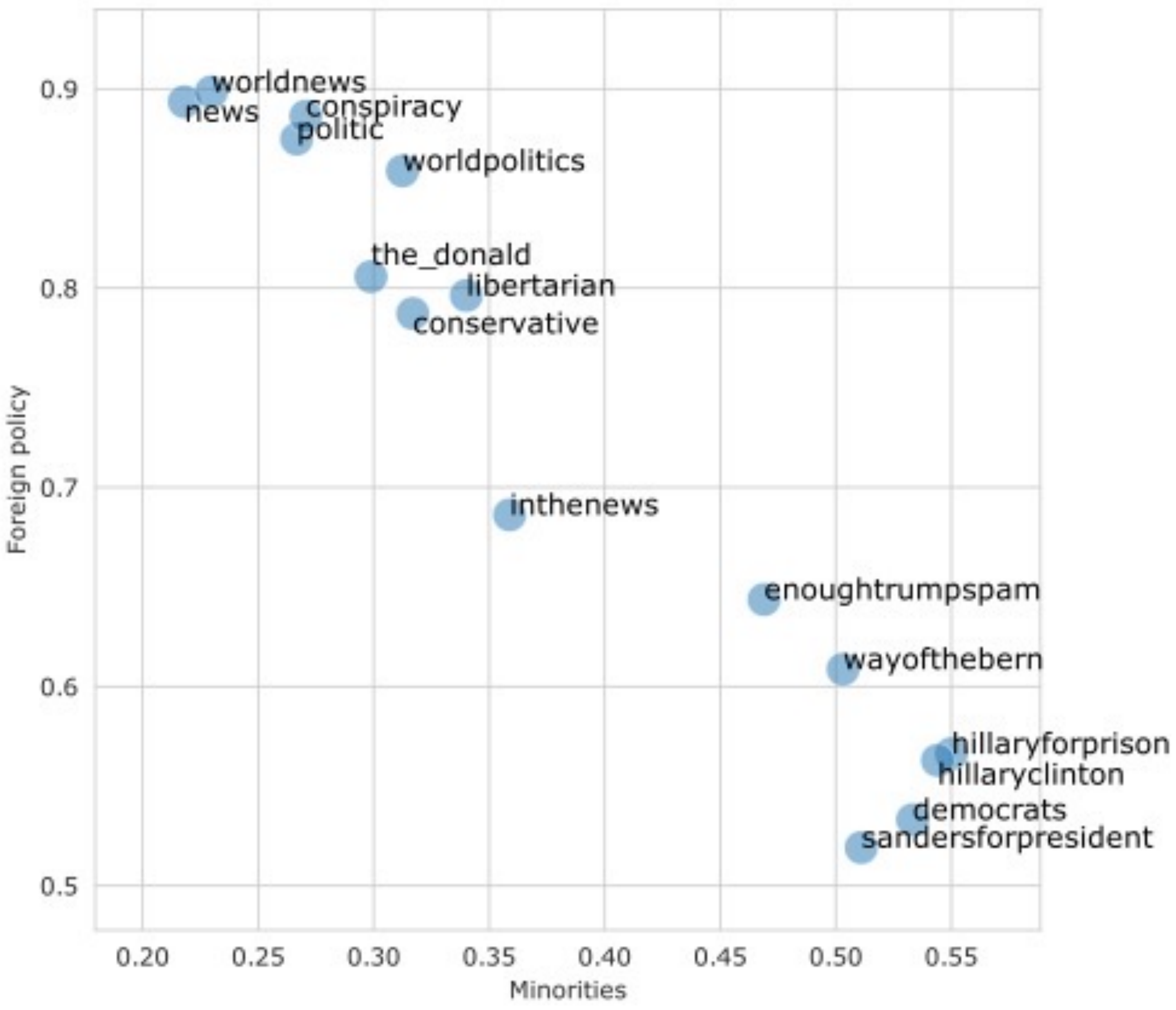
            end for
        end for
    end for
end for
Return ϕ and θ.

Results
Real data & Interpretability

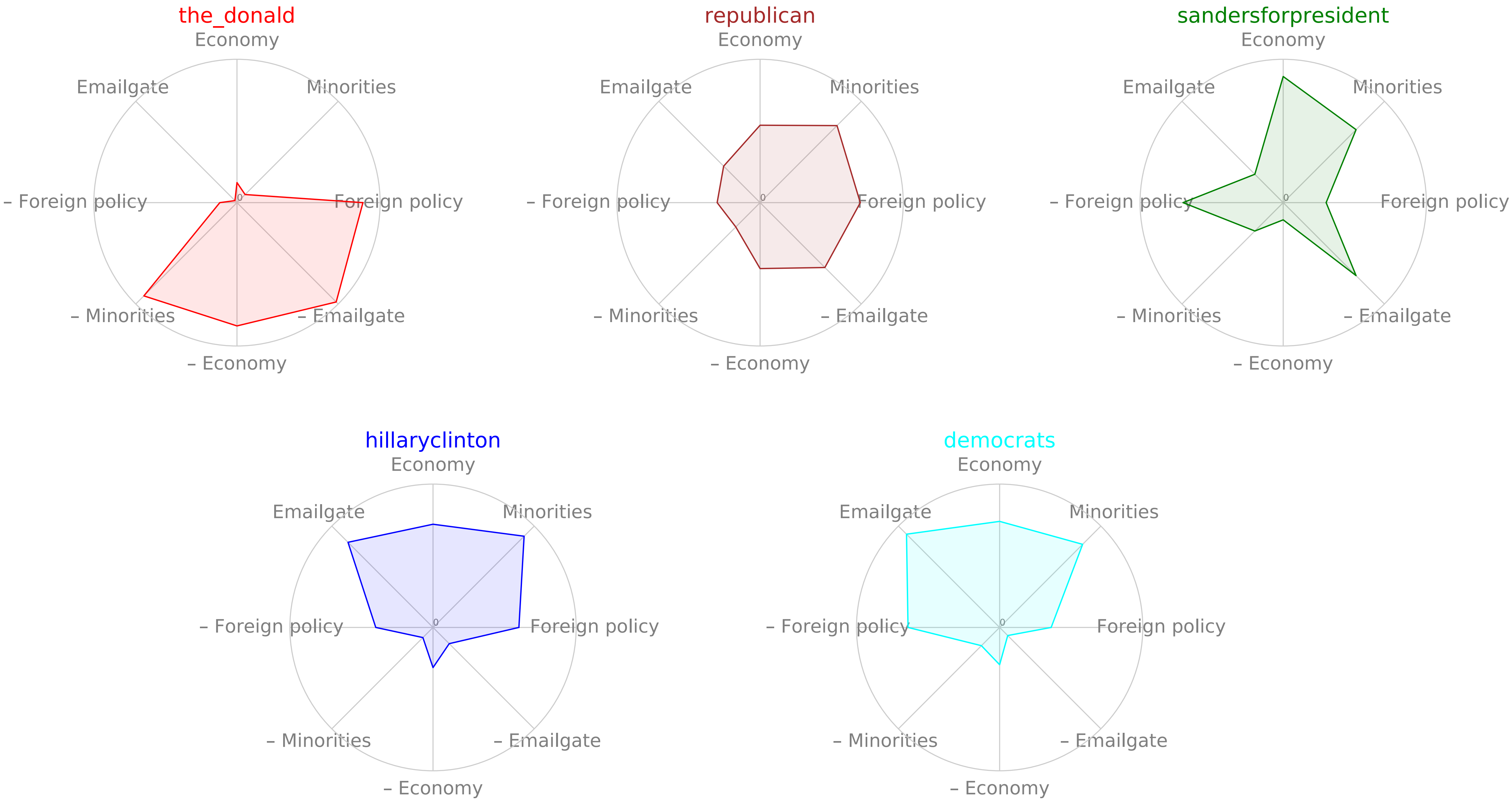
- Reddit
  - Nodes: 50 political subreddits
  - Actions: Posting a Url
    - 22.047 items
  - 5 topics identified
    - Economy, emailgate, Foreign policy, Campaign, Minorities



democrats
hillaryclinton
enoughtrumpspam
inthenews
sandersforpresident
worldnews
libertarian
worldpolitics
politic
news
wayofthebern
conservative
the_donald
hillaryforprison
conspiracy
0.55
0.50
0.45
0.40
0.35
0.30
Emailgate
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
Economy



worldnews
news
conspiracy
politic
worldpolitics
the_donald
libertarian
conservative
inthenews
enoughtrumpspam
wayofthebern
hillaryforprison
hillaryclinton
democrats
sandersforpresident
0.9
0.8
0.7
0.6
0.5
Foreign policy
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
Minorities



the_donald
Economy
Minorities
Foreign policy
– Emailgate
– Economy
– Minorities
– Foreign policy
Emailgate
0
republican
Economy
Minorities
Foreign policy
– Emailgate
– Economy
– Minorities
– Foreign policy
Emailgate
0
sandersforpresident
Economy
Minorities
Foreign policy
– Emailgate
– Economy
– Minorities
– Foreign policy
Emailgate
0
hillaryclinton
Economy
Minorities
Foreign policy
– Emailgate
– Economy
– Minorities
– Foreign policy
Emailgate
0
democrats
Economy
Minorities
Foreign policy
– Emailgate
– Economy
– Minorities
– Foreign policy
Emailgate
0

Conclusions

- Representation learning for modeling information propagation
  - Susceptibility and influence
  - Interest and relevance
  - Speed of the propagation process
  - Polarity and alignments
- Advantages
  - Interpretability, predictive power, scalable learning
- A unified perspective
  - Can we put all these pieces together?
- Why is it important?
  - Echo-chamber detection
  - Moderating discussions, mitigation policies
  - Fighiting misinformation

Thank you

Questions?

giuseppe.manco@icar.cnr.it

@beman