Learning Ideological Embeddings from Information Cascades

Corrado Monti, Francesco Bonchi

ISI Foundation & ISI Global Science Foundation

Giuseppe Manco,

Istituto di Calcolo e Reti ad Alte Prestazioni

Cigdem Aslay,

Aalto University

Why Ideological Embeddings?

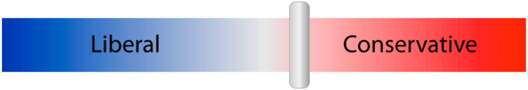
- Many works in opinion modeling assume one axis

Why Ideological Embeddings?

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

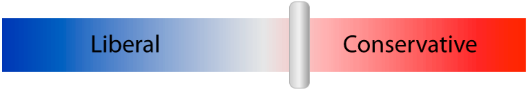
Why Ideological Embeddings?

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



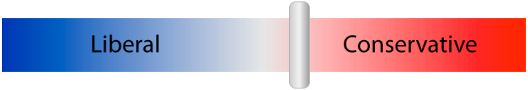
Why Ideological Embeddings?

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



Why Ideological Embeddings?

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

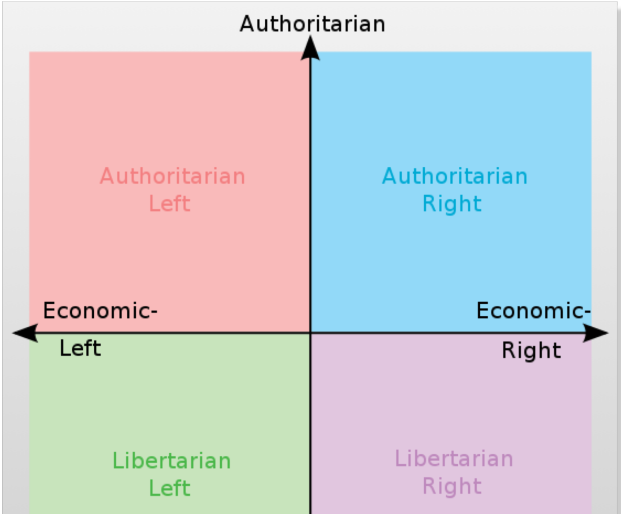


Why Ideological Embeddings?

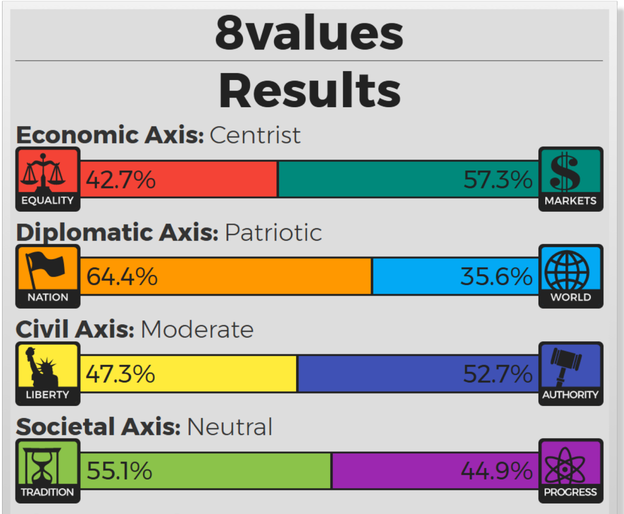
Idea: a multi-dimensional ideological space

Why Ideological Embeddings?

Idea: a multi-dimensional ideological space

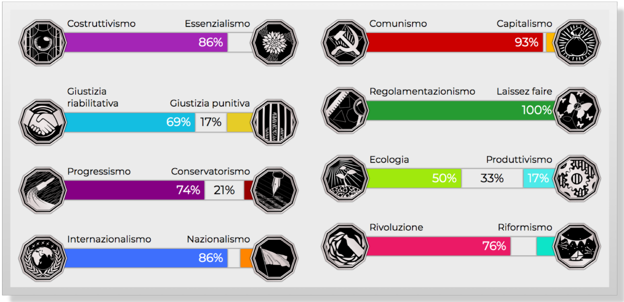


Idea: a multi-dimensional ideological space



Why Ideological Embeddings?

Idea: a multi-dimensional ideological space



Why Ideological Embeddings?

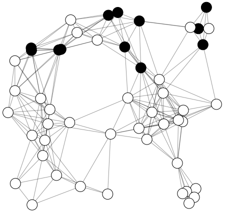
Idea: a multi-dimensional ideological space

- Which behavior might make this observable?

Why Ideological Embeddings?

Idea: a multi-dimensional ideological space

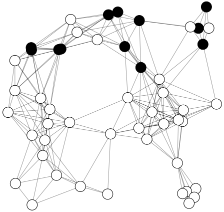
- Which behavior might make this observable?
- Spreading news on social media



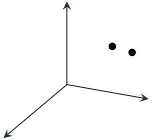
Why Ideological Embeddings?

Idea: a multi-dimensional ideological space

- Which behavior might make 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*



More in general

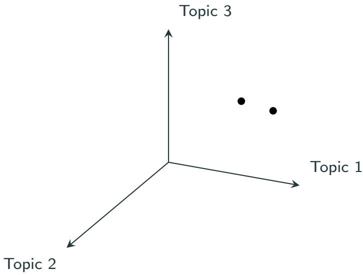
More in general

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

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

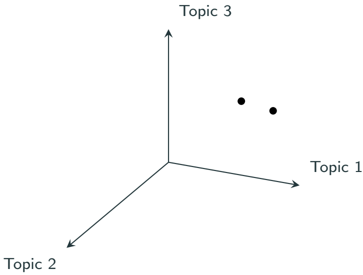
How?

- How to build such a space?
- Let any researcher *define their axis*



How?

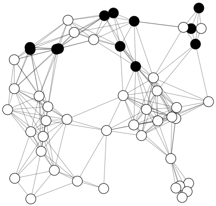
- How to build such a space?
- Let any researcher *define their axis*
- As input, each item must be tagged
    - a news can be “healthcare’ or “migration”
    - Manually or LDA
    - Can be fuzzy
- Those define the axes of our space



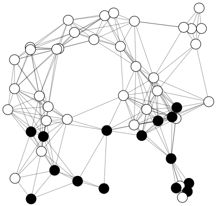
An example

○ = user, ● = user sharing an item

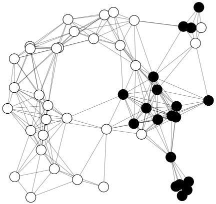
Item 1
0.9 healthcare +
0.1 migration



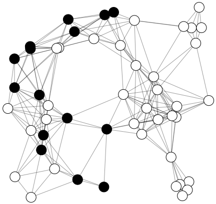
Item 2
0.9 healthcare +
0.1 migration



Item 3
0.1 healthcare +
0.9 migration



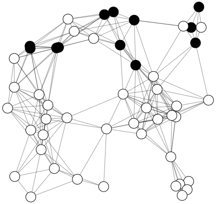
Item 4
0.1 healthcare +
0.9 migration



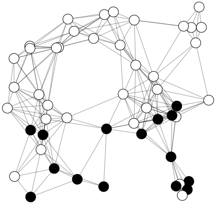
An example

○ = user, ● = user sharing an item

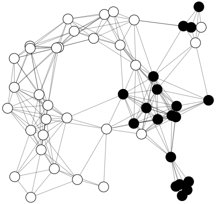
Item 1
0.9 healthcare +
0.1 migration



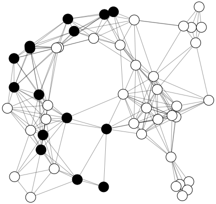
Item 2
0.9 healthcare +
0.1 migration



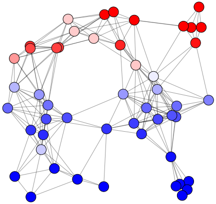
Item 3
0.1 healthcare +
0.9 migration



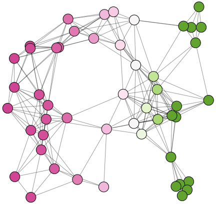
Item 4
0.1 healthcare +
0.9 migration



healthcare

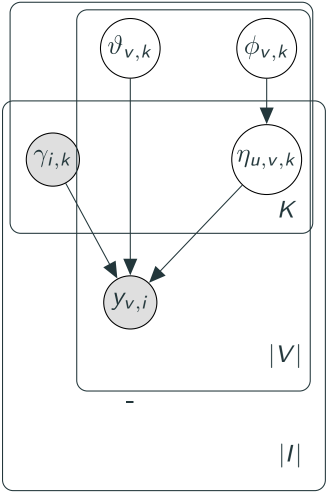


migration



Formalize these assumptions into a model

The model



- $\vartheta_{v,z} \in [0,1] \rightarrow$ node interests
- $\phi_{v,z} \in [0,1] \rightarrow$ node polarities
- $\eta_{u,v,z} \in [0,1] \rightarrow$ polarity alignment
- $\gamma_{i,z} \in [0,1] \rightarrow$ item topics
- $y \in \{0,1\} \rightarrow$ activation

The algorithm

- Likelihood of observed item i is $\Pr(\mathcal{D}_i|\Theta) =$

$$\prod_{u \in \mathcal{D}_i} \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}) \cdot \prod_{u \notin \mathcal{D}_i} (1 - \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}))$$

  where $F_{i,u}$ are potential activators

The algorithm

- Likelihood of observed item i is $\Pr(\mathcal{D}_i|\Theta) =$

$$\prod_{u \in \mathcal{D}_i} \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}) \cdot \prod_{u \notin \mathcal{D}_i} (1 - \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}))$$

  where $F_{i,u}$ are potential activators

- Simplifying assumption:
  $1 - \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}) \approx \prod_{v \in F_{i,u}} (1 - \Pr(u \in \mathcal{D}_i|\Theta, v \in \mathcal{D}_i))$,
  since we are looking for clusters of alignment

The algorithm

- Likelihood of observed item i is $\Pr(\mathcal{D}_i|\Theta) =$

$$\prod_{u \in \mathcal{D}_i} \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}) \cdot \prod_{u \notin \mathcal{D}_i} (1 - \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}))$$

  where $F_{i,u}$ are potential activators

- Simplifying assumption:
  $1 - \Pr(u \in \mathcal{D}_i|\Theta, F_{i,u}) \approx \prod_{v \in F_{i,u}} (1 - \Pr(u \in \mathcal{D}_i|\Theta, v \in \mathcal{D}_i))$,
  since we are looking for clusters of alignment

- In this way, we obtain a scalable OGD algorithm:
  - Each potential activation $u \to v$ is an example
  - Positive or negative if succeeded or not

The algorithm

Input: $G = (V, E)$; items $i \in \mathcal{I}$ with topics $\gamma_i \in \mathbb{R}^K$, activations $\mathcal{D}_i \subseteq V$.

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:

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

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

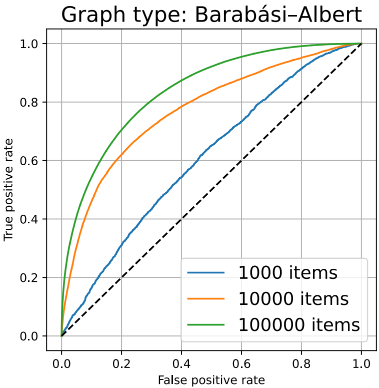
$$\nabla_{\phi,\vartheta} \log \left(1 - \sum_k \gamma_{i,k} \cdot \vartheta_{u,k} \cdot p(u, v, k) \right)$$

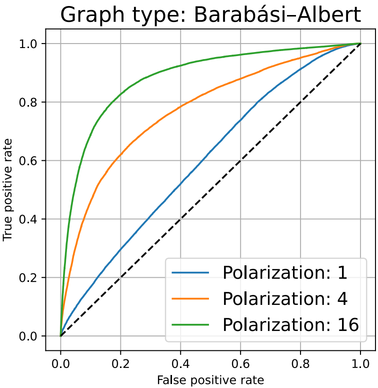
Results: synthetic data

Results on synthetic data

Num. items	AUC ROC	Avg. Prec.
1000	0.607	0.543
10000	0.773	0.733
100000	0.840	0.815

p	AUC ROC	Avg. Prec.
1	0.601	0.534
4	0.773	0.733
16	0.884	0.857





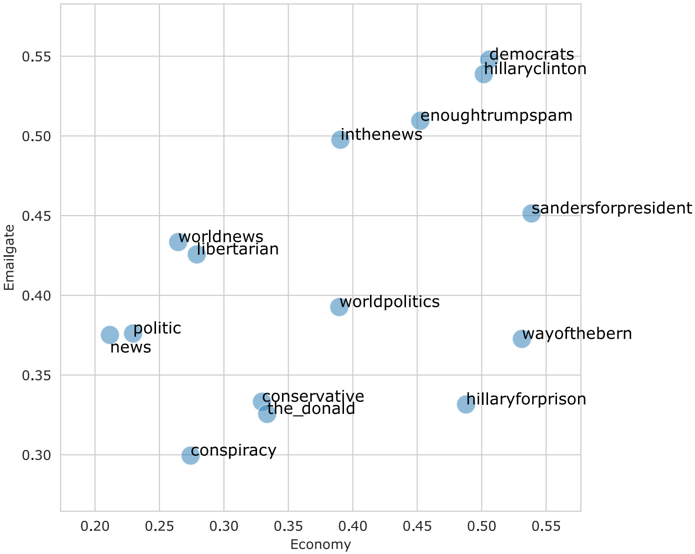
Results: real data & interpretability

Reddit Dataset

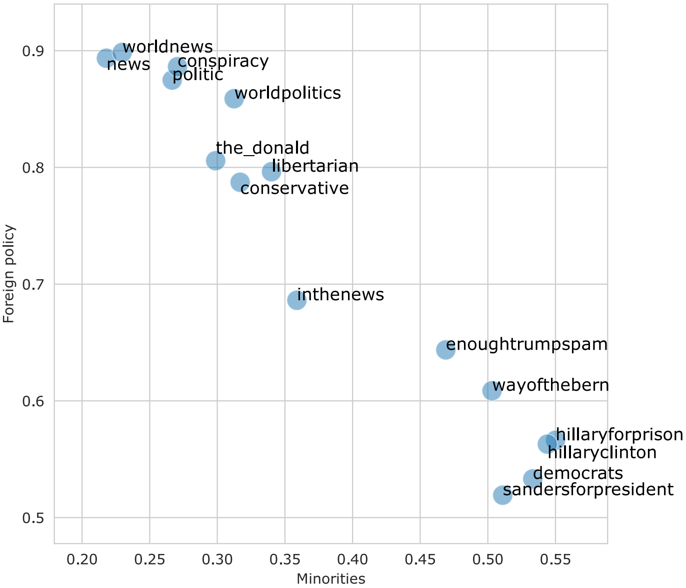
- Nodes: 50 political subreddits
- Action: posting a URL $\rightarrow$ 22 047 items
- Graph: complete
- 5 topics identified automatically with `doc2vec`
  - Economy, Emailgate, Foreign policy, Campaign, Minorities

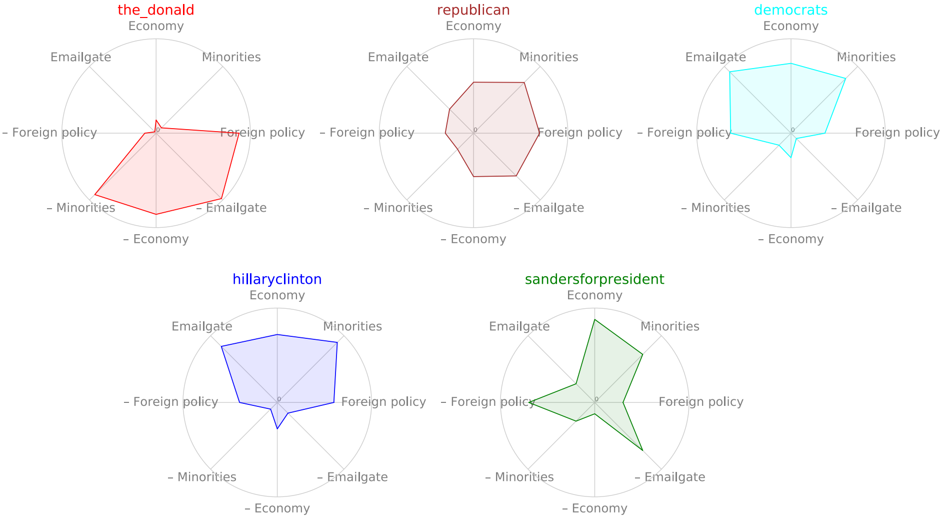
Follows assumptions:

action spreads if nodes are ideologically aligned



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

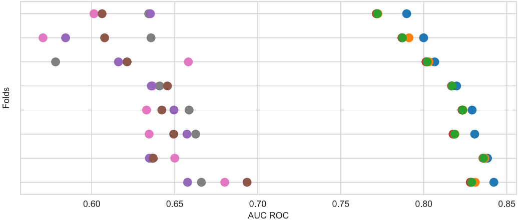

worldnews
news
conspiracy
politic
worldpolitics
the_donald
libertarian
conservative
inthenews
enoughtrumpspam
wayofthebern
hillaryforprison
hillaryclinton
democrats
sandersforpresident
0.9
0.8
0.7
0.6
0.5
Foreign policy
0.20
0.25
0.30
0.35
0.40
0.45
0.50
0.55
Minorities


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

Results: real data & predictive power

Results: real data & predictive power

Reddit dataset:



	Algorithm	AUC ROC	Difference	Avg. Precision	Time (s)
●	Our MIP model	0.820 ± 0.019	0.000 ± 0.000	0.777 ± 0.022	7.1
●	Barbera model	0.812 ± 0.022	-0.008 ± 0.005	0.790 ± 0.018	6.4
●	Original inf. + Topics	0.811 ± 0.022	-0.009 ± 0.006	0.781 ± 0.018	10.1
●	Original information	0.810 ± 0.022	-0.009 ± 0.006	0.785 ± 0.018	9.5
●	node2vec, d=128	0.639 ± 0.028	-0.180 ± 0.024	0.660 ± 0.033	11.6
●	node2vec + Topics, d=128	0.638 ± 0.028	-0.182 ± 0.016	0.650 ± 0.031	14.5
●	node2vec + Topics, d=11	0.634 ± 0.024	-0.186 ± 0.019	0.634 ± 0.033	8.7
●	node2vec, d=11	0.633 ± 0.034	-0.186 ± 0.024	0.641 ± 0.036	7.3

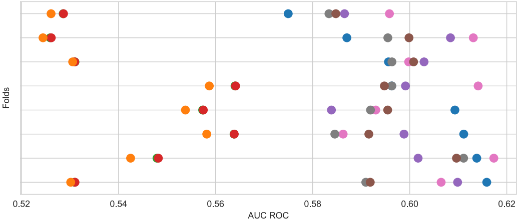
Results: real data & predictive power

Twitter dataset:

- Nodes: 738 users
- Items: 3 624 retweets
- Topics: political hashtags
- Graph: retweet

Results: real data & predictive power

Twitter dataset:



	Algorithm	AUC ROC	Difference	Avg. Precision	Time (s)
●	Our MIP model	0.601 ± 0.015	0.000 ± 0.000	0.435 ± 0.013	83.1
●	node2vec, d=11	0.603 ± 0.011	0.003 ± 0.018	0.431 ± 0.010	480.5
●	node2vec + Topics, d=128	0.596 ± 0.007	-0.004 ± 0.014	0.427 ± 0.008	561.7
●	node2vec + Topics, d=11	0.599 ± 0.009	-0.002 ± 0.015	0.425 ± 0.009	485.3
●	node2vec, d=128	0.594 ± 0.009	-0.007 ± 0.014	0.425 ± 0.006	523.1
●	Original inf. + Topics	0.544 ± 0.016	-0.057 ± 0.016	0.391 ± 0.020	985.7
●	Original information	0.544 ± 0.016	-0.057 ± 0.016	0.391 ± 0.020	965.4
●	Barbera model	0.541 ± 0.015	-0.060 ± 0.015	0.387 ± 0.019	75.9

Conclusions

Conclusions

- **Ideological embeddings**
  - Multidimensional, axes defined through items
  - Interpretable
- Our model is
  - an **interpretation** of reality
  - a way to make **predictions**
- Algorithm combines interpretability, predictive power and scalability
- Probabilistic model can be valid tools for social science

Code & data:

https://github.com/corradomonti/ideological-embeddings

Thanks!

corrado.monti@isi.it