POLE: Polarized Embedding for Signed Networks

WSDM'22

Zexi Huang[1], Arlei Silva[2], Ambuj Singh[1]

[1]Department of Computer Science, University of California, Santa Barbara

[2]Department of Computer Science, Rice University

Research funded by NSF IIS 1817046 and HDTRA1-19-1-0017.

Social polarization

Donald J. Trump @realDonaldTrump — Following

TODAY WE MAKE AMERICA GREAT AGAIN!

Donald J. Trump @realDonaldTrump — Follow

The Fake News is working overtime. Just reported that, despite the tremendous success we are having with the economy & all things else, 91% of the Network News about me is negative (Fake). Why do we work so hard in working with the media when it is corrupt? Take away credentials?

Donald J. Trump @realDonaldTrump

These are the things and events that happen when a sacred landslide election victory is so unceremoniously & viciously stripped away from great patriots who have been badly & unfairly treated for so long. Go home with love & in peace. Remember this day forever!

This claim of election fraud is disputed, and this Tweet can't be replied to, Retweeted, or liked due to a risk of violence

@

Replying to @realDonaldTrump

These damned fake news outlets need to be taken down! If they had done this to Obama, it would've been the end of the world!

We are with you, Mr President!

@

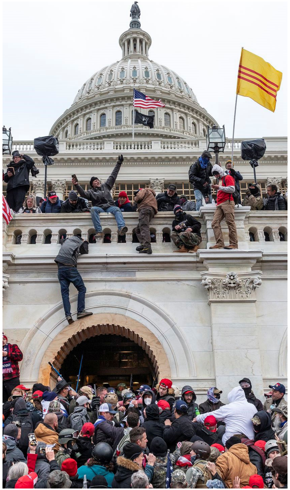
Replying to @realDonaldTrump

The [D] party will cease to exist once it's all exposed.
FAKE NEWS can no longer control [dampen] public awareness of the TRUTH.
DARK TO LIGHT.
Q

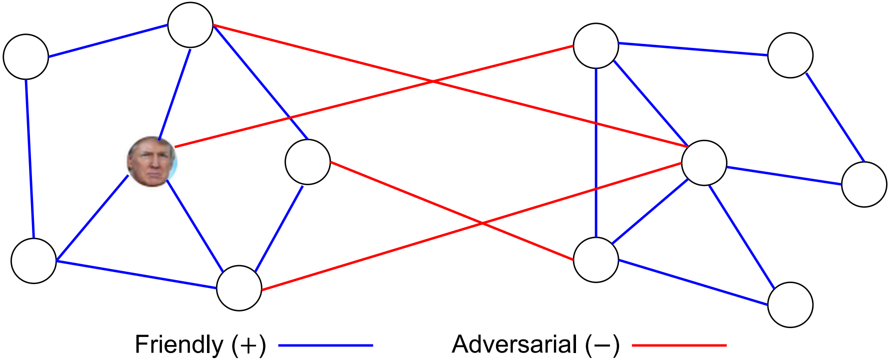
@

Replying to @realDonaldTrump

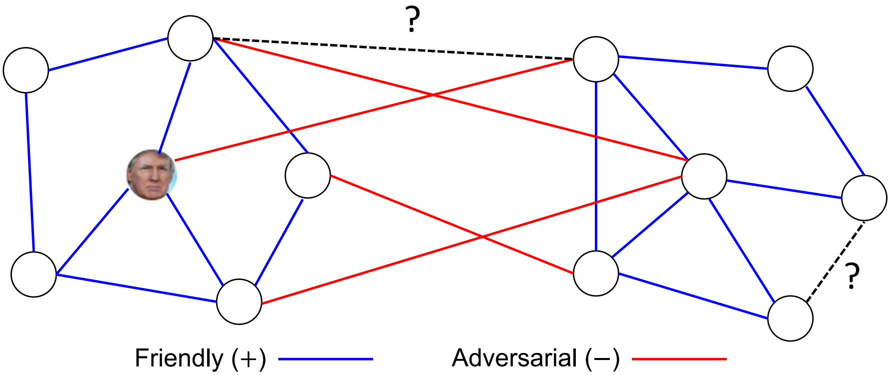
I love seeing Trump supporters CRY, it's my daily medicine, my weekly energy, my monthly inspiration and my yearly motivation. Their loss is the only reason i'm still alive, i was born to love and enjoy the failure that they have achieved.



Signed networks



Signed networks



Can we predict future conflicts in signed networks to reduce further polarization?

Signed link prediction in polarized networks

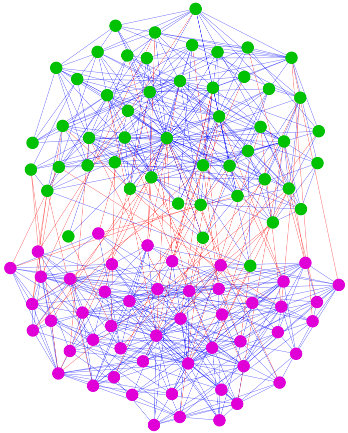
- Predicting signs of links:
  - Signed embedding [1, 2] (signed similarity)

Signed link prediction in polarized networks

- Predicting signs of links:
  - Signed embedding [1, 2] (signed similarity)
- What about predicting link existence?
  - Unsigned embedding [3, 4] (connectivity)

Signed link prediction in polarized networks

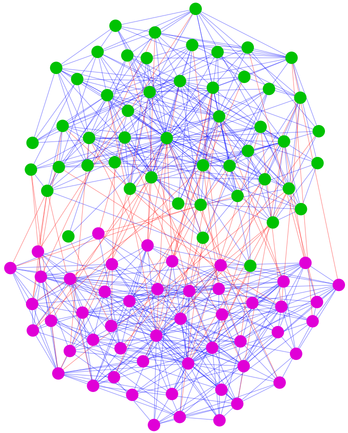
- Predicting signs of links:
  - Signed embedding [1, 2] (signed similarity)
- What about predicting link existence?
  - Unsigned embedding [3, 4] (connectivity)



Intra-community links: dense, positive
Inter-community links: sparse, negative

Signed link prediction in polarized networks

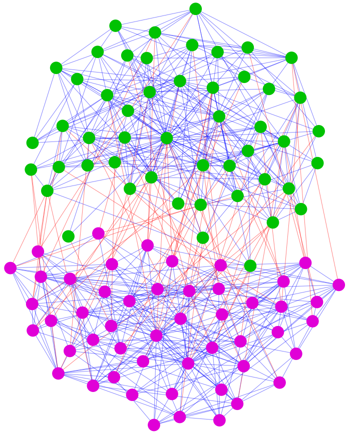
- Predicting signs of links:
  - Signed embedding [1, 2] (signed similarity)
- What about predicting link existence?
  - Unsigned embedding [3, 4] (connectivity)
- Unable to predict negative links between communities



Intra-community links: dense, positive
Inter-community links: sparse, negative

Signed link prediction in polarized networks

- Predicting signs of links:
  - Signed embedding [1, 2] (signed similarity)
- What about predicting link existence?
  - Unsigned embedding [3, 4] (connectivity)
- Unable to predict negative links between communities
  - Need to capture signed/unsigned similarity jointly



Intra-community links: dense, positive
Inter-community links: sparse, negative

Signed random walk

- Unsigned RW: $|M|_{uv}(t) = \sum_{\text{all length-}t\text{ paths }l\text{ between }u\text{ and }v} \text{Prob}(l)$
  - $\text{Prob}(l)$ captures the unsigned similarity.

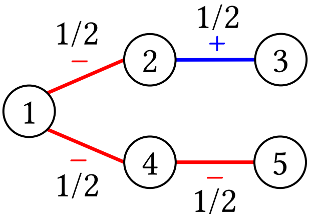
Signed random walk

- Unsigned RW: $|M|_{uv}(t) = \sum_{\text{all length-}t\text{ paths }l\text{ between }u\text{ and }v} \text{Prob}(l)$
  - $\text{Prob}(l)$ captures the unsigned similarity.

- Signed RW: $M_{uv}(t) = \sum_{\text{all length-}t\text{ paths }l\text{ between }u\text{ and }v} \text{Prob}(l)\text{Sign}(l)$
  - $\text{Sign}(l)$ based on the social balance theory captures the signed similarity.

Signed random walk

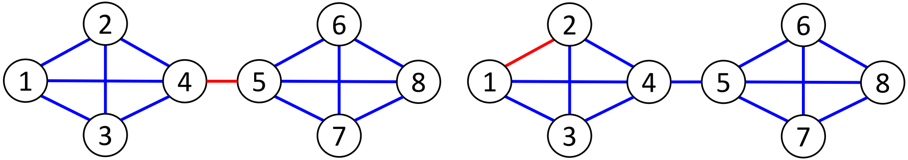
- Unsigned RW: $|M|_{uv}(t) = \sum_{\text{all length-}t\text{ paths }l\text{ between }u\text{ and }v} \text{Prob}(l)$
  - $\text{Prob}(l)$ captures the unsigned similarity.
- Signed RW: $M_{uv}(t) = \sum_{\text{all length-}t\text{ paths }l\text{ between }u\text{ and }v} \text{Prob}(l)\text{Sign}(l)$
  - $\text{Sign}(l)$ based on the social balance theory captures the signed similarity.



l	$\text{Prob}(l)$	$\text{Sign}(l)$
$\langle 1, 2, 3 \rangle$	$1/2 \times 1/2 = 1/4$	$-1 \times 1 = -1$
$\langle 1, 4, 5 \rangle$	$1/2 \times 1/2 = 1/4$	$-1 \times -1 = 1$

Measuring polarization

- $|M|_{:u}(t)$ and $M_{:u}(t)$ are highly correlated if u is polarized.



(a) (b)

$$|M|_{:1} = [0.23, 0.21, 0.21, 0.18, 0.05, 0.02, 0.02, 0.02]$$

$$M^a_{:1} = [0.23, 0.21, 0.21, 0.18, -0.05, -0.02, -0.02, -0.02]$$

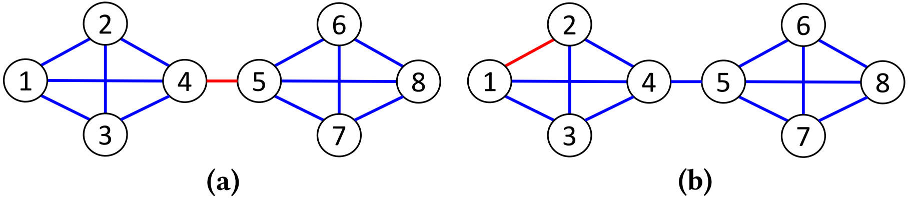
$$M^b_{:1} = [0.12, -0.01, 0.09, 0.07, 0.03, 0.01, 0.01, 0.01]$$

$$\mathrm{corr}(|M|_{:1}, M^a_{:1}) = 0.9849$$

$$\mathrm{corr}\left(|M|_{:1}, M^b_{:1}\right) = 0.6237$$

Measuring polarization

- $|M|_{:u}(t)$ and $M_{:u}(t)$ are highly correlated if u is polarized.



$$|M|_{:1} = [0.23, 0.21, 0.21, 0.18, 0.05, 0.02, 0.02, 0.02]$$

$$M^a_{:1} = [0.23, 0.21, 0.21, 0.18, -0.05, -0.02, -0.02, -0.02]$$

$$M^b_{:1} = [0.12, -0.01, 0.09, 0.07, 0.03, 0.01, 0.01, 0.01]$$

$$\text{corr}(|M|_{:1}, M^a_{:1}) = 0.9849$$

$$\text{corr}\left(|M|_{:1}, M^b_{:1}\right) = 0.6237$$

- $\text{Pol}(u;t) = \text{corr}(|M|_{:u}(t), M_{:u}(t))$, $\text{Pol}(G;t) = \text{mean}_{u \in G}(\text{Pol}(u;t))$.

POLE: polarized embedding

- POLE: extends autocovariance similarity [5, 6] to signed RWs

$$R(t) = M(t)^T W M(t)$$

where $W = \frac{1}{\mathrm{vol}(G)} D - \frac{1}{\mathrm{vol}(G)^2} d d^T$.

POLE: polarized embedding

- POLE: extends autocovariance similarity [5, 6] to signed RWs

$$R(t) = M(t)^T W M(t)$$

where $W = \frac{1}{\mathrm{vol}(G)} D - \frac{1}{\mathrm{vol}(G)^2} dd^T$.

- Desired properties:
  - Positive links: large positive similarity;
  - Negative links: large negative similarity;
  - Non-links: small similarity.
- Predict negative inter-community links as most dissimilar pairs.

Comparison of similarity distributions

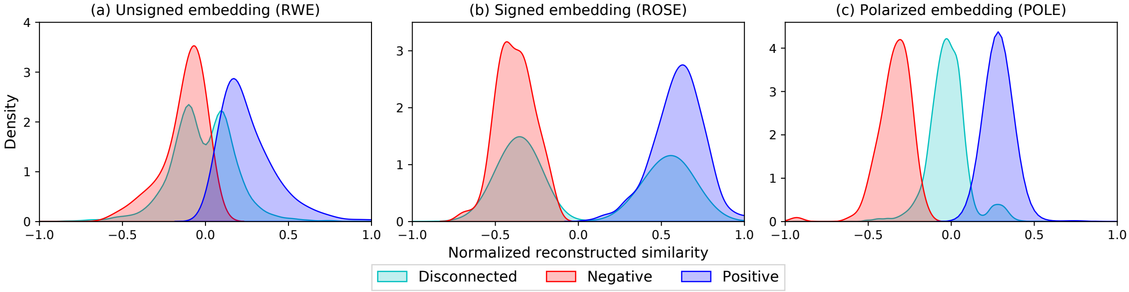


Figure: Distributions of the reconstructed similarity for different types of node pairs in a polarized network using (a) unsigned [6], (b) signed [7], and (c) polarized embedding.

Comparison of similarity distributions

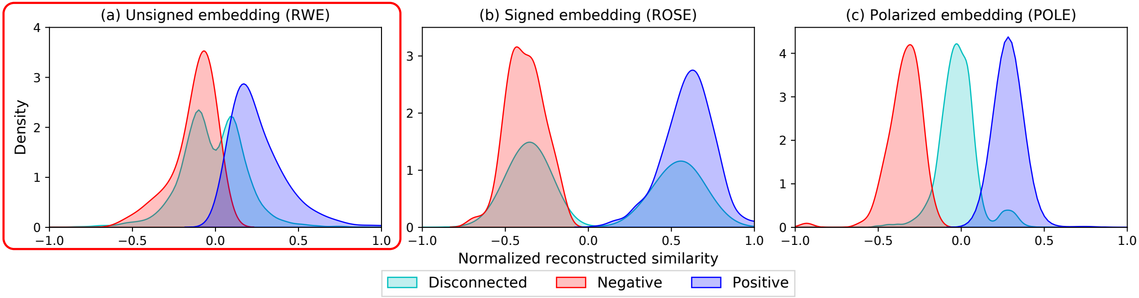


Figure: Distributions of the reconstructed similarity for different types of node pairs in a polarized network using (a) unsigned [6], (b) signed [7], and (c) polarized embedding.

Comparison of similarity distributions

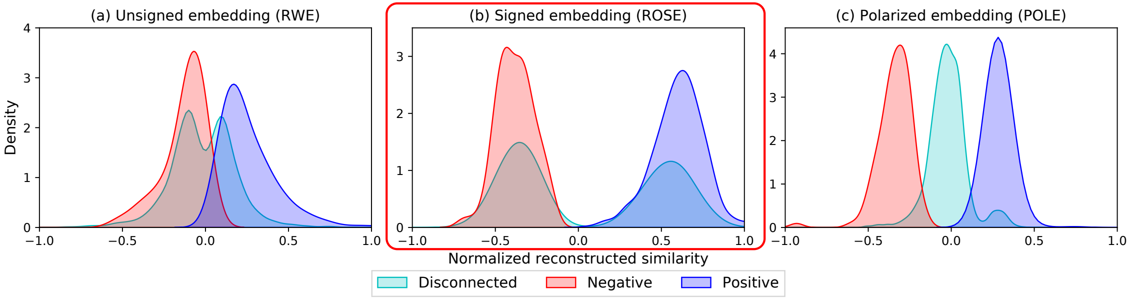


Figure: Distributions of the reconstructed similarity for different types of node pairs in a polarized network using (a) unsigned [6], (b) signed [7], and (c) polarized embedding.

Comparison of similarity distributions

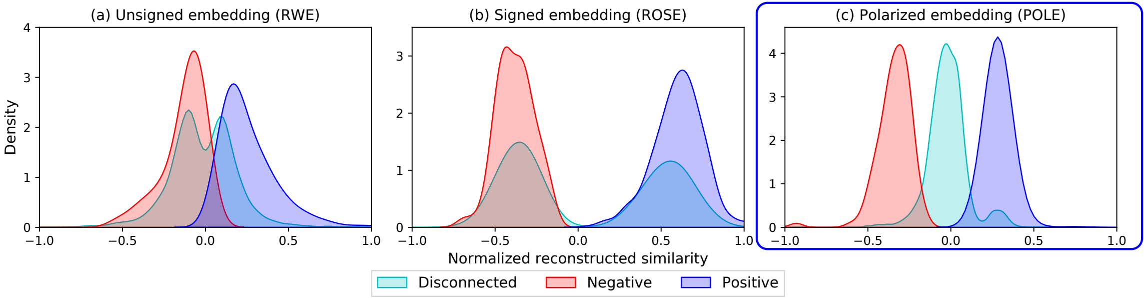


Figure: Distributions of the reconstructed similarity for different types of node pairs in a polarized network using (a) unsigned [6], (b) signed [7], and (c) polarized embedding.

Datasets:

	$\lvert\mathcal{V}\rvert$	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{E}_-\rvert/\lvert\mathcal{E}\rvert$
Congress	219	523	20.46%
WoW-EP8	789	116,009	18.63%
Bitcoin-Alpha	3,772	14,077	9.31%
Bitcoin-OTC	5,872	21,431	14.71%
Referendum	10,864	251,396	5.09%
Wiki-RfA	11,275	169,925	22.04%

Baselines: SiNE [8], SIGNet [9], SIDE [1], BESIDE [2], SLF [10], ROSE [7]

Signed link prediction setting:

- POLE: compute dot product similarity
- Baselines: train two classifiers (positive/negative vs non-links)
- Evaluation metric: positive/negative precision@k

Signed link prediction

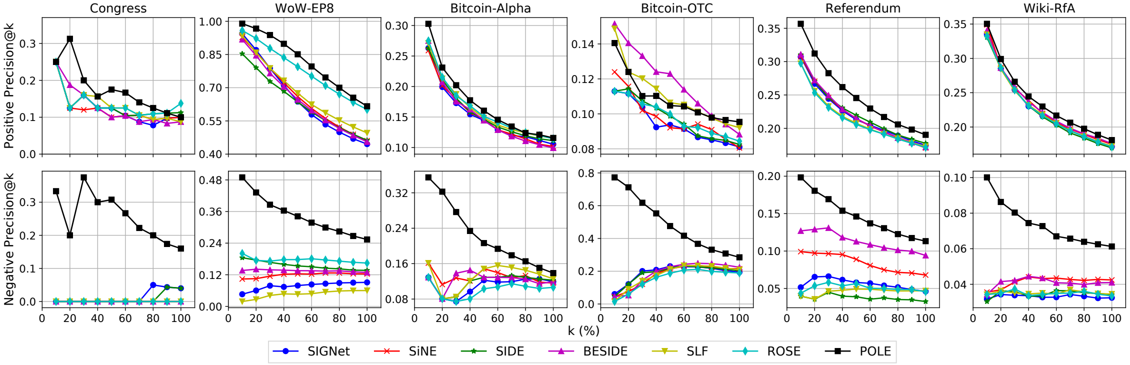


Figure: Signed link prediction performance comparison between POLE and baselines. POLE outperforms all baselines in almost all datasets, especially for the negative links.

Measuring polarization

Table: Ten least polarized congresspeople by our RW-based polarization measure.

Congressperson	State	Party	Score
Henry Cuellar*	Texas	D	-0.6542
Jane Harman†	California	D	-0.5376
Curt Weldon	Pennsylvania	R	-0.4381
Dutch Ruppersberger	Maryland	D	-0.4318
Jim Moran	Virginia	D	-0.3832
Dave Obey	Wisconsin	D	-0.3588
Wayne Gilchrest	Maryland	R	-0.3503
Duke Cunningham	California	R	-0.3248
Al Edwards	Texas	D	-0.3063
Lincoln Davis	Tennessee	D	-0.2901

* "Voted with President Trump 75% of time" — 538
† "Best Republican in the Democratic Party" — LA Times

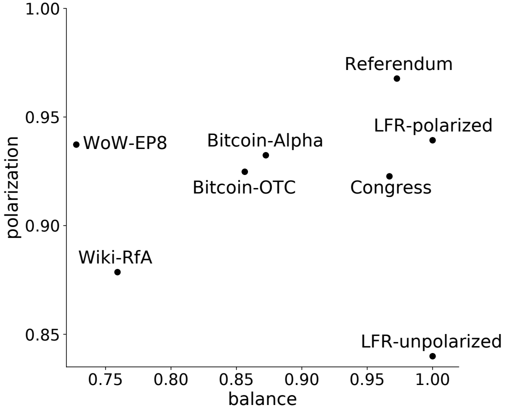
Measuring polarization

Table: Ten least polarized congresspeople by our RW-based polarization measure.

Congressperson	State	Party	Score
Henry Cuellar*	Texas	D	-0.6542
Jane Harman†	California	D	-0.5376
Curt Weldon	Pennsylvania	R	-0.4381
Dutch Ruppersberger	Maryland	D	-0.4318
Jim Moran	Virginia	D	-0.3832
Dave Obey	Wisconsin	D	-0.3588
Wayne Gilchrest	Maryland	R	-0.3503
Duke Cunningham	California	R	-0.3248
Al Edwards	Texas	D	-0.3063
Lincoln Davis	Tennessee	D	-0.2901

* "Voted with President Trump 75% of time" — 538
† "Best Republican in the Democratic Party" — LA Times

Figure: Polarization and social balance of real-world graphs and LFR benchmarks.



This work investigated the representation learning problem for signed link prediction in polarized networks, where both links and their signs have to be inferred jointly.

Contributions:

- A characterization of the challenges of negative link prediction.
- A random-walk based polarization measure for signed networks.
- A novel signed embedding model (POLE) for polarized networks.

References I

[1] Junghwan Kim, Haekyu Park, Ji-Eun Lee, and U Kang.
Side: representation learning in signed directed networks.
In *WebConf*, 2018.

[2] Yiqi Chen, Tieyun Qian, Huan Liu, and Ke Sun.
"bridge" enhanced signed directed network embedding.
In *CIKM*, 2018.

[3] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena.
Deepwalk: Online learning of social representations.
In *SIGKDD*, 2014.

[4] Aditya Grover and Jure Leskovec.
node2vec: Scalable feature learning for networks.
In *SIGKDD*, 2016.

[5] Michael T Schaub, Jean-Charles Delvenne, Renaud Lambiotte, and Mauricio Barahona.
Multiscale dynamical embeddings of complex networks.
PRE, 99(6):062308, 2019.

References II

[6] Zexi Huang, Arlei Silva, and Ambuj Singh.
A broader picture of random-walk based graph embedding.
In *SIGKDD*, 2021.

[7] Amin Javari, Tyler Derr, Pouya Esmailian, Jiliang Tang, and Kevin Chen-Chuan Chang.
Rose: Role-based signed network embedding.
In *WebConf*, 2020.

[8] Suhang Wang, Jiliang Tang, Charu Aggarwal, Yi Chang, and Huan Liu.
Signed network embedding in social media.
In *SDM*, 2017.

[9] Mohammad Raihanul Islam, B Aditya Prakash, and Naren Ramakrishnan.
Signet: Scalable embeddings for signed networks.
In *PAKDD*, 2018.

[10] Pinghua Xu, Wenbin Hu, Jia Wu, and Bo Du.
Link prediction with signed latent factors in signed social networks.
In *SIGKDD*, 2019.