ROSE: Role-based Signed Network Embedding

Amin Javari[1] Tyler Derr[2] Pouya Esmailian[3] Jiliang Tang[2]
Kevin Chen-Chuan Chang[1]

[1]University of Illinois at Urbana-Champaign

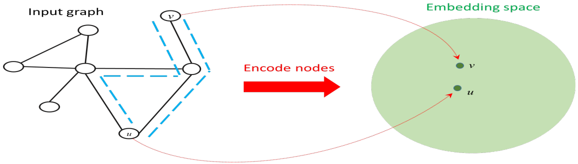
[2]Michigan State University

[3]Sharif University of Technology

April 22, 2020

Node Embedding

Goal: Map nodes to an embedding space in way that their similarity in the original network can be approximated based on their similarity in the embedding space.



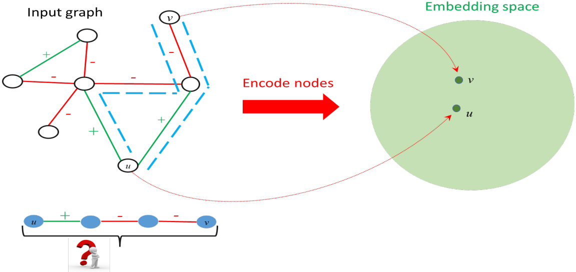
Four components of an embedding model[1]:

- A pairwise node similarity function
  - A path between two nodes indicates their closeness.
- An encoder function
- A decoding function
- A loss function

[1] William Hamilton et al. "Representation learning on graphs: Methods and applications". In: *IEEE Data Engineering* (2017).

Node Embedding in Signed Networks

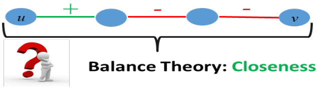
Signed Network is defined as $G(V, E)$ with a link type mapping function $\varphi : E \to A$ where for each link $e \in E$, $\varphi(e) \in A$ and $A = \{+, -\}$.



The unsigned similarity functions cannot be directly applied to signed networks because paths containing negative edges do not necessarily represent closeness.

Why Node Embedding in Signed Networks is Challenging?

- Existing methods embed nodes closer if the path between them indicates closeness, and distance them otherwise.[2]
- To interpret if a path indicates closeness or distantness, some strong assumptions are exploited.
  - **Balance theory:** a cycle is balanced if there exist an even number of negative links[3].



- The strong assumptions naturally induce noise to the embedding process.
- This strategy does not use a principled way to distant nodes based on the absence of paths between them.

[2] Yiqi Chen et al. "Bridge: Enhanced Signed Directed Network Embedding". In: *CIKM*. 2018.

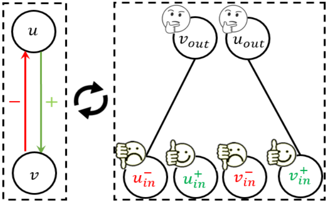
[3] Junghwan Kim et al. "Side: Representation learning in signed directed networks". In: *WWW*. 2018.

Main Idea: Network Transformation based Embedding

The input network can be transformed to another network in which we do not encounter the embedding challenges present in the original network.

1. Network transformation
2. Embedding the transformed network
3. Embedding the original network by aggregating the embeddings of the transformed network.

ROSE relies on the notion of transformation based embedding.

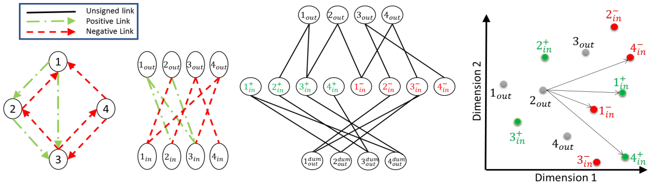


ROSE: Network Transformation and Embedding the Transformed Network

Transformation to the network of role-nodes: Define different roles for a node, denoted as **role-nodes**.

1. Transformation to a bipartite network
2. Transformation to an unsigned network
3. Augmenting the network.

Embedding the network of role-nodes: A classic embedding model can be used to embed role-nodes, e.g., node2vec.



(a) Input signed network (b) Bipartited signed network (c) Bipartited unsigned network of role-nodes (d) Embedding space of role-nodes

ROSE: Embedding the Original Network

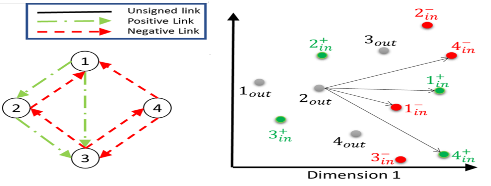
A node's embedding is created by aggregating the embeddings of the corresponding role-nodes.

Aggregation methods

- Fixed aggregation, e.g., concatenation: $W_u = W_{u_{out}} || W_{u_{in}^+} || W_{u_{in}^-}$.
- Target aware aggregation
  - Inspired by recommender systems, out role-node of u can be embedded according to v.
  - The target dependent embedding of u w.r.t. v is defined as $W_u^v = W_{u_{out}^v} || W_u$.
  - $W_{u_{out}^v}$ can be obtained by attending to the neighbors of u_{out} based on their relevancy to v.

ROSE: Model Justification

- ROSE does not rely on any assumption about the origin of the network.

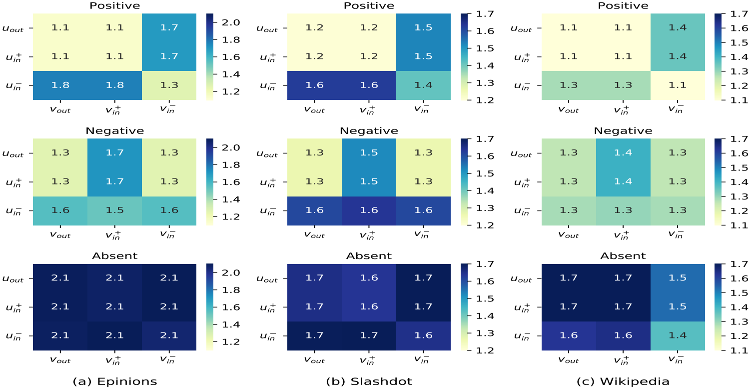


Input signed network

Embedding space of role-nodes

- The model preserves both link labels and link structure.
  - If a link with label l exists from u to v, $W_{u_{out}}$ has higher proximity to $W_{v^l_{in}}$ than $W_{v^{l'}_{in}}$.
  - If there is no link from from u to v, $W_{u_{out}}$ is expected to have low proximities to both $W_{v^l_{in}}$ and $W_{v^{l'}_{in}}$.
- Flexibility and generalizability.

Experiments: Interpretation of the encodings of role-nodes



The average pairwise distance of the encoding vectors of the role-nodes of a node pair (u, v) for different interaction-types between them.

The distance values are consistent with the introduced patterns.

Experiments: Performance of the proposed model

AUC of the proposed models and the baseline methods.

Model	Sign Prediction			Link Prediction		
	WikiElection	**Slashdot**	**Epinions**	**WikiElection**	**Slashdot**	**Epinions**
SIDE[4]	0.7986	0.8815	0.8672	0.9184	0.9342	0.9314
BESIDE[5]	0.8953	0.9012	0.9342	0.9092	0.9265	0.9397
SiNE[6]	0.8632	0.8680	0.8543	0.5833	0.5983	0.6488
SIGNET[7]	0.8943	0.8997	0.9181	0.9099	0.8862	0.9205
ROSE	0.9091	0.9082	0.9533	0.9418	0.9357	0.9403
ROSE-UAT	**0.9116**	**0.9095**	**0.9547**	**0.9426**	**0.9391**	**0.9444**

- ROSE has superior performance than the baseline models.
- ROSE-UAT perform better than ROSE. Encoding the nodes with respect to a target entity helps to better analyze their interactions.

[4] Junghwan Kim et al. "Side: Representation learning in signed directed networks". In: *WWW*. 2018.
[5] Yiqi Chen et al. "Bridge: Enhanced Signed Directed Network Embedding". In: *CIKM*. 2018.
[6] Suhang Wang et al. "Signed network embedding in social media". In: *SIAM*. 2017.
[7] Mohammad Islam et al. "Signet: Scalable embeddings for signed networks". In: *PKDD*. 2018.

Thank you!