Signed Bipartite Graph Neural Networks

Junjie Huang[1,3], Huawei Shen[1,3], Qi Cao[1], Shuchang Tao[1,3], Xueqi Cheng[2]

[1]Data Intelligence System Research Center,
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2]CAS Key Laboratory of Network Data Science and Technology,
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[3]University of Chinese Academy of Sciences, Beijing, China
{huangjunjie17s, shenhuawei, caoqi, taoshuchang18z, cxq}@ict.ac.cn

September 28, 2021

Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

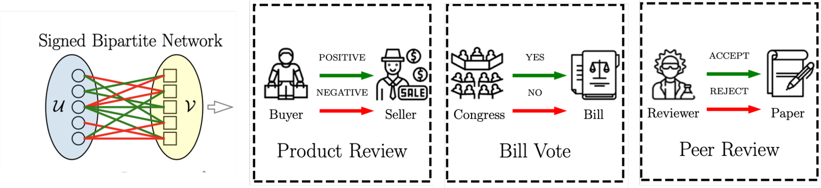
Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

Signed Bipartite Networks

- Consider a signed bipartite network, $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$, where $\mathcal{U} = \{u_1, u_2, ..., u_{|\mathcal{U}|}\}$ and $\mathcal{V} = \{v_1, v_2, ..., v_{|\mathcal{V}|}\}$ represent two sets of nodes with the number of nodes $|\mathcal{U}|$ and $|\mathcal{V}|$. $\mathcal{E} \subset \mathcal{U} \times \mathcal{V}$ is the edges between $\mathcal{U}$ and $\mathcal{V}$. $\mathcal{E} = \mathcal{E}^+ \bigcup \mathcal{E}^-$ is the set of edges between the two sets of nodes $\mathcal{U}$ and $\mathcal{V}$ where $\mathcal{E}^+ \bigcap \mathcal{E}^- = \varnothing$, $\mathcal{E}^+$ and $\mathcal{E}^-$ represent the sets of positive and negative edges, respectively.

It is commonly found in many fields including business, politics, and academics, but has been less studied.



Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

Signed Graph Modeling

Signed networks are such social networks having both positive and negative links.

Balance theory is the fundamental theory in the signed network field.

- For classical signed networks, signed triangles are the most common way to measure the balance of signed networks.

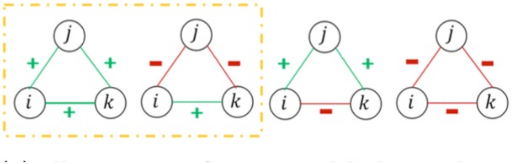


Figure: Illustration of structural balance theory. (Huang et al., ICANN2019)

Graph representation learning

- Matrix factorization-based methods
  - HOPE (Ou et al., 2016, KDD2016)
- Random-walk based algorithms
  - DeepWalk (Perozzi et al., KDD2014), Node2vec (Grover and Leskovec, KDD2016), LINE (Tang et al., WWW2015)
  - BiNE (Gao et al., SIGIR2016)
- Graph neural networks
  - GCN (Kipf and Welling, ICLR2017), GraphSAGE (Hamilton et al., NIPS2017), GAT (Veličković et al., ICLR2018), GIN (Xu et al., ICLR2018)

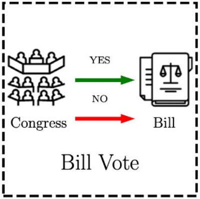
Signed graph representation learning

- Signed network embeddings
  - SiNE (Wang et al., SIAM2017), SIDE (Kim et al., WWW2018)
- Signed GNNs
  - SGCN (Derr et al., ICDM2018), SDGNN (Huang et al., AAAI2021)

Outline

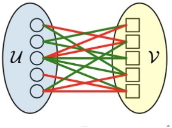
1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

In different scenarios, the negative ratio varies.



	Bonanza	U.S. House	U.S. Senate	Preliminary Review	Final Review
$\|\mathcal{U}\|$	7,919	515	145	182	182
$\|\mathcal{V}\|$	1,973	1,281	1,056	304	304
$\|\mathcal{E}\| = \|\mathcal{E}^+\| + \|\mathcal{E}^-\|$	36,543	114,378	27,083	1,170	1,170
% Positive Links	0.980	0.540	0.553	0.403	0.397
% Negative Links	0.020	0.460	0.447	0.597	0.603

Signed Bipartite Network



Signed Caterpillars, Signed Butterflies and Signed Triangles

Two different perspectives:

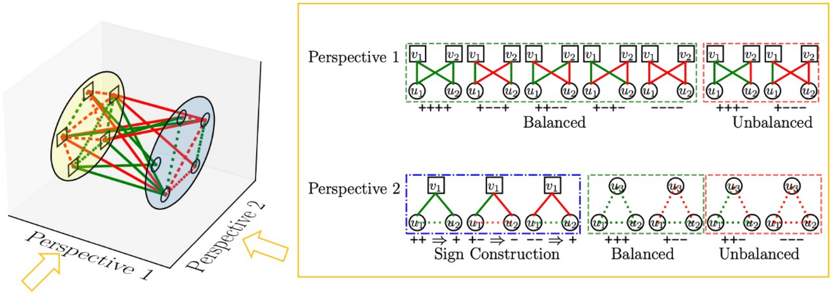


Figure: Perspective 1 offers to analyze the signed butterfly isomorphism. For Perspective 2, we can analyze the signed triangle isomorphism by sign construction.

Findings:

- Large majority of signed butterflies/signed triangles in signed bipartite networks are more balanced than expectation.
- In the scenario of peer reviews, after rebuttal phase, the balance of signed bipartite networks increased.

	Bonanza	U.S. House	U.S. Senate	Preliminary Review	Final Review
Signed Butterfly Isomorphism ++++ (%, %E)	(0.986, 0.922)	(0.244, 0.085)	(0.262, 0.094)	(0.109, 0.026)	(0.115↑, 0.025)
Signed Butterfly Isomorphism +−−+ (%, %E)	(0.000, 0.001)	(0.109, 0.123)	(0.108, 0.122)	(0.109, 0.116)	(0.072↓, 0.115)
Signed Butterfly Isomorphism ++−− (%, %E)	(0.001, 0.001)	(0.111, 0.123)	(0.110, 0.122)	(0.101, 0.116)	(0.057↓, 0.115)
Signed Butterfly Isomorphism +−+− (%, %E)	(0.000, 0.001)	(0.186, 0.123)	(0.184, 0.122)	(0.156, 0.116)	(0.215↑, 0.115)
Signed Butterfly Isomorphism −−−− (%, %E)	(0.000, 0.000)	(0.147, 0.045)	(0.133, 0.040)	(0.249, 0.127)	(0.315↑, 0.133)
Balanced Signed Butterfly Summary (%, %E)	(**0.988**, 0.924)	(**0.798**, 0.500)	(**0.798**, 0.500)	(**0.724**, 0.501)	(**0.774↑**, 0.501)
Signed Butterfly Isomorphism +++− (%, %E)	(0.012, 0.076)	(0.118, 0.289)	(0.122, 0.302)	(0.070, 0.156)	(0.075↑, 0.151)
Signed Butterfly Isomorphism +−−− (%, %E)	(0.000, 0.000)	(0.085, 0.211)	(0.081, 0.197)	(0.206, 0.343)	(0.151↓, 0.349)
Unbalanced Signed Butterfly Summary (%, %E)	(**0.012**, 0.076)	(**0.202**, 0.500)	(**0.202**, 0.500)	(**0.276**, 0.499)	(**0.226↓**, 0.499)
Signed Triangles Isomorphism +++ in $\mathcal{U}$ (%, %E)	(0.978, 0.949)	(0.338, 0.217)	(0.360, 0.248)	(0.327, 0.213)	(0.446↑, 0.310)
Signed Triangles Isomorphism +−− in $\mathcal{U}$ (%, %E)	(0.011, 0.001)	(0.476, 0.287)	(0.436, 0.261)	(0.451, 0.290)	(0.346↓, 0.212)
Balanced Signed Triangles Summary in $\mathcal{U}$ (%, %E)	(**0.989**, 0.950)	(**0.815**, 0.504)	(**0.796**, 0.508)	(**0.778**, 0.504)	(**0.792↑**, 0.522)
Signed Triangle Isomorphism ++− in $\mathcal{U}$ (%, %E)	(0.011, 0.050)	(0.176, 0.432)	(0.189, 0.440)	(0.194, 0.431)	(0.195↑, 0.444)
Signed Triangle Isomorphism −−− in $\mathcal{U}$ (%, %E)	(0.000, 0.000)	(0.009, 0.063)	(0.015, 0.051)	(0.027, 0.065)	(0.012↓, 0.034)
Unbalanced Signed Triangles Summary in $\mathcal{U}$(%, %E)	(**0.011**, 0.050)	(**0.185**, 0.496)	(**0.204**, 0.492)	(**0.222**, 0.496)	(**0.208↓**, 0.478)

Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

Message/Aggregation/Update Function

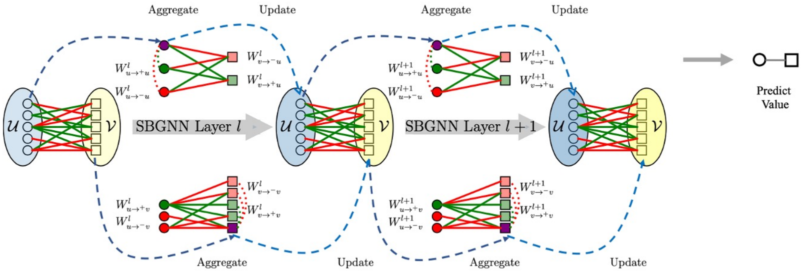


Figure: Illustration of SBGNN. SBGNN Layer includes Aggeregate and Update functions. The aggregated message comes from the Set_1 and Set_2 with positive and negative links. After getting the embedding of the node u_i and v_i, it can be used to predict the link sign relationship.

GNNs Message Passing Scheme

$$m_{j\to i}^{(l)}(i,j) = \mathrm{MSG}^{(l)}\left(h_i^{(l)}, h_j^{(l)}, h_{e_{ji}}^{(l)}\right), j \in \mathcal{N}(j),$$

$$m_{j\to i}^{(l)}(i) = \mathrm{AGG}^{(l)}\left(\left\{m_{j\to i}^{(l)}(i,j) \mid j \in \mathcal{N}(i)\right\}\right),$$

$$h_i^{(l+1)} = \mathrm{UPT}^{(l)}\left(h_i^{(l)}, m_{j\to i}^{(l)}(i)\right),$$

Loss Function

- After getting embeddings $z_{u_i} \in \mathbb{R}^{d_u}$ and $z_{v_j} \in \mathbb{R}^{d_v}$ of the node u_i and v_j, we can use following methods to get the prediction value for $u_i \to v_j$.
  - product operation:

    $$y_{pred} = \mathrm{sigmoid}(z_{u_i}^\top \cdot z_{v_j}),$$

    where $\cdot^\top$ is the transpose function and $\mathrm{sigmoid}$ is the sigmoid function $f(x) = \frac{1}{1+e^{-x}}$.
  - $\textsc{Mlp}$:

    $$y_{pred} = \mathrm{sigmoid}\big(\textsc{Mlp}(z_{u_i} \parallel z_{v_i})\big)$$

    where $\textsc{Mlp}$ is a two layer neural networks, $\parallel$ is the concatenation operation.
- After getting the prediction values, we use binary cross entropy as the loss function:

$$\mathcal{L} = -w\left[y \cdot \log y_{pred} + (1-y) \cdot \log(1-y_{pred})\right]$$

where w is the rescaling weight for the unblanced negative ratios; y is the ground truth with mapping $\{-1, 1\}$ to $\{0, 1\}$.

Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

Experimental Settings

- Datasets
  - Bonanza, Review, U.S. House, U.S. Senate
- Tasks: LINK SIGN PREDICTION (binary classification problem)
- Train/Val/Test: 85/5/10 (5 times train/val/test splits)
- Baselines:
  - Random Embeddings
  - Unsigned Network Embeddings: DeepWalk, Node2vec, LINE
  - Signed/Bipartite Network Embedding: SiNE, BiNE, SBiNE
  - Signed Butterfly Based Methods: SCsc, MFwBT, SBRW
  - Signed Bipartite Graph Neural Networks: SBGNN-MEAN, SBGNN-GAT
- Evaluation Metrics
  - AUC, Binary-F1, Macro-F1, and Micro-F1

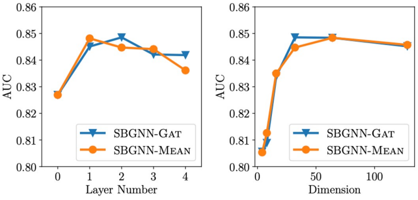
Experiment results

Table: The results of Link Sign Prediction on four datasets.

		Random Embedding	Unsigned Network Embedding			Signed/Bipartite Network Embedding			Signed Butterfly Based Methods			Signed Bipartite Graph Neural Networks	
Dataset	Metric	Random	Deepwalk	Node2vec	LINE	SiNE	BiNE	SBiNE	SCsc	MFwBT	SBRW	SBGNN-MEAN	SBGNN-GAT
Bonanza	AUC	0.5222	0.6176	0.6185	0.6124	0.6088	0.6026	0.5525	**0.6524**	0.5769	0.5315	0.5841	0.5769
	Binary-F1	0.7282	0.7843	0.7530	0.6974	0.9557	0.7426	0.8514	0.6439	0.8927	**0.9823**	0.9488*	0.9616*
	Macro-F1	0.3868	0.4258	0.4087	0.3790	**0.5422**	0.4016	0.4538	0.3543	0.4813	0.5353	0.5311*	0.5404*
	Micro-F1	0.5770	0.6497	0.6093	0.5424	0.9157	0.5960	0.7436	0.4843	0.8076	**0.9652**	0.9044*	0.9269*
Review	AUC	0.5489	0.6324	0.6472	0.6236	0.5741	#N/A	0.5329	0.5522	0.4727	0.5837	0.6584*	**0.6747***
	Binary-F1	0.4996	0.5932	0.6141	0.5974	0.5247	#N/A	0.4232	0.3361	0.4346	0.5423	0.6128*	**0.6366***
	Macro-F1	0.5426	0.6268	0.6400	0.6120	0.5688	#N/A	0.5262	0.4823	0.4696	0.5767	0.6556*	**0.6629***
	Micro-F1	0.5487	0.6325	0.6444	0.6137	0.5744	#N/A	0.5521	0.5812	0.4752	0.5812	0.6632*	**0.6667***
U.S. House	AUC	0.5245	0.6223	0.6168	0.5892	0.6006	0.6103	0.8328	0.8274	0.8097	0.8224	0.8474*	**0.8481***
	Binary-F1	0.5431	0.6401	0.6323	0.6304	0.6118	0.6068	0.8434	0.8375	0.8234	0.8335	0.8549*	**0.8560***
	Macro-F1	0.5238	0.6215	0.6158	0.5883	0.5991	0.6097	0.8323	0.8267	0.8096	0.8219	0.8463*	**0.8471***
	Micro-F1	0.5246	0.6224	0.6166	0.5892	0.5996	0.6108	0.8330	0.8274	0.8106	0.8226	0.8468*	**0.8476***
U.S. Senate	AUC	0.5251	0.6334	0.6260	0.5743	0.5875	0.6071	0.7998	0.8163	0.7857	0.8142	0.8209*	**0.8246***
	Binary-F1	0.5502	0.6603	0.6526	0.6159	0.5923	0.5968	0.8175	0.8294	0.8043	0.8291	0.8277	**0.8320**
	Macro-F1	0.5239	0.6325	0.6251	0.5722	0.5842	0.6037	0.7992	0.8148	0.7850	0.8131	0.8177*	**0.8215***
	Micro-F1	0.5254	0.6347	0.6271	0.5732	0.5848	0.6042	0.8009	0.8160	0.7867	0.8145	0.8183*	**0.8221***

- Modeling the balance theory in the signed bipartite network is key for LINK SIGN PREDICTION .
- Our SBGNN models outperform other baseline models.

Parameter Analysis and Ablation Study



(a) #Layer l **(b) Dimension d**

Figure: Parameter analysis on the number of SBGNN Layer l and imension d for SBGNN on the U.S. House dataset.

Table: Ablation study results for SBGNN model on the U.S. House dataset.

Method	AUC	Binary-F1	Macro-F1	Micro-F1
SBGNN-GAT	0.8485	0.8586	0.8477	0.8485
SBGNN-GAT (w/o Set_1)	0.8406	0.8521	0.8400	0.8409
SBGNN-GAT (w/o Set_2)	0.8440	0.8567	0.8438	0.8448
SBGNN-GAT (with LR)	0.6281	0.6195	0.6227	0.6227
SBGNN-GAT (with MLP)	0.8365	0.8480	0.8358	0.8367
SBGNN-MEAN	0.8447	0.8519	0.8429	0.8434
SBGNN-MEAN (w/o Set_1)	0.8419	0.8496	0.8402	0.8408
SBGNN-MEAN (w/o Set_2)	0.8296	0.8410	0.8288	0.8297
SBGNN-MEAN (with LR)	0.6285	0.6387	0.6263	0.6267
SBGNN-MEAN (with MLP)	0.8443	0.8531	0.8430	0.8436

- d=32 and l=2 can have a better result.
- MLP is much better than simple LR but not better than product operation.

Outline

1. Background
2. Related Work
3. Balance Theory in Signed Bipartite Networks
4. Proposed Methodology
5. Experiments
6. Conclusions and Future Work

Conclusions and Future Work

Conclusions

- We discuss two different perspectives to model the signed bipartite networks.
- We further use these two perspectives to model peer review and find that after rebuttal, the balance of reviewers' opinions improved.
- Under the definition of a new perspective, we propose a new graph neural network model SBGNN to learn the node representation of signed bipartite graphs .
- Our SBGNN model achieve the state-of-the art performance in several datasets.

Future work

- We will explore signed bipartite networks with node features.
- We will try to introduce signed bipartite graph neural networks into recommender system.

Reference I

T. Derr, Y. Ma, and J. Tang. Signed graph convolutional networks. In *2018 IEEE International Conference on Data Mining (ICDM)*, pages 929–934. IEEE, 2018.

M. Gao, L. Chen, X. He, and A. Zhou. Bine: Bipartite network embedding. In *The 41st international ACM SIGIR conference on research & development in information retrieval*, pages 715–724, 2018.

A. Grover and J. Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016.

W. L. Hamilton, R. Ying, and J. Leskovec. Inductive representation learning on large graphs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 1025–1035, 2017.

J. Huang, H. Shen, L. Hou, and X. Cheng. Signed graph attention networks. In *International Conference on Artificial Neural Networks*, pages 566–577. Springer, 2019.

J. Huang, H. Shen, L. Hou, and X. Cheng. Sdgnn: Learning node representation for signed directed networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 196–203, 2021.

J. Kim, H. Park, J.-E. Lee, and U. Kang. Side: representation learning in signed directed networks. In *Proceedings of the 2018 World Wide Web Conference*, pages 509–518, 2018.

T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations*, 2019.

M. Ou, P. Cui, J. Pei, Z. Zhang, and W. Zhu. Asymmetric transitivity preserving graph embedding. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1105–1114, 2016.

B. Perozzi, R. Al-Rfou, and S. Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710, 2014.

J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei. Line: Large-scale information network embedding. In *Proceedings of the 24th international conference on world wide web*, pages 1067–1077, 2015.

P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio. Graph attention networks. *International Conference on Learning Representations*, 2018. URL https://openreview.net/forum?id=rJXMpikCZ.

S. Wang, J. Tang, C. Aggarwal, Y. Chang, and H. Liu. Signed network embedding in social media. In *Proceedings of the 2017 SIAM international conference on data mining*, pages 327–335. SIAM, 2017.

K. Xu, W. Hu, J. Leskovec, and S. Jegelka. How powerful are graph neural networks? In *International Conference on Learning Representations*, 2018.