Signed Graph Neural Network with Latent Groups

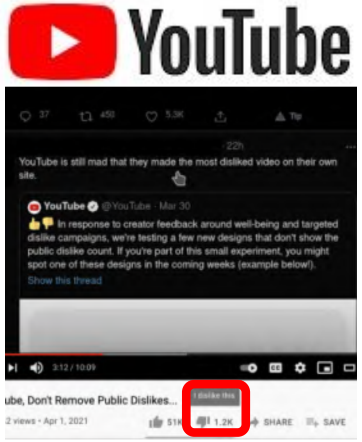
Haoxin Liu*	Ziwei Zhang*	Peng Cui	Yafeng Zhang	Qiang Cui	Jiashuo Liu	Wenwu Zhu
Tsinghua U	Tsinghua U	Tsinghua U	Meituan	Meituan	Tsinghua U	Tsinghua U

* Equal contribution

Besides Positive Relationships

Negative relationships also play an important role

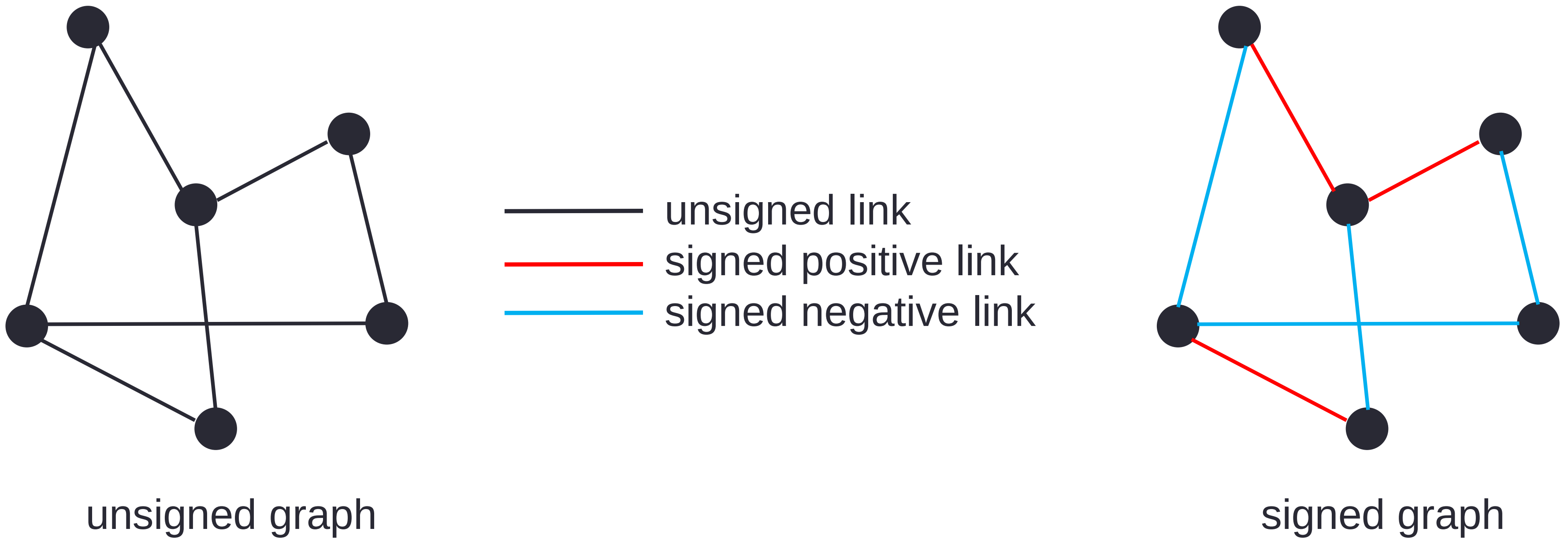
- Foes
- Disagreements
- Boycotts
- Dislike
- Distrust
- ……



How to model positive and negative relationships **simultaneously**?

Modeling: Signed Graph

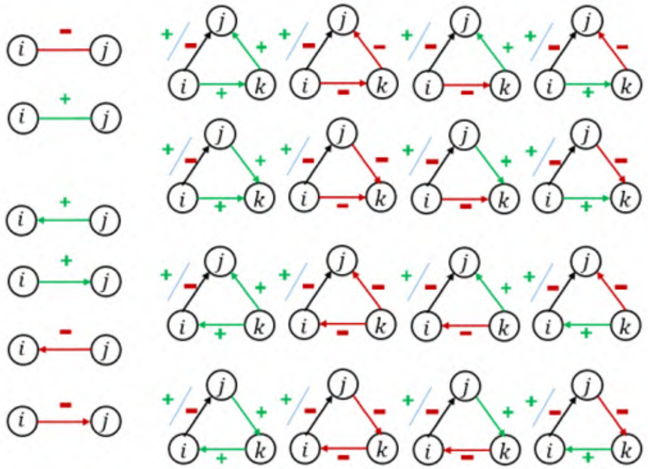
Assign signs to links



Challenges

Unsinged graph representation learning methods not suitable

- Fundamental hypotheses in unsinged graphs fail
  - e.g. homophily
- Complex semantic relationships
  - e.g. $(2^{k+1}-1)$ types of relationships considering k-order neighbors
  - e.g. 38 popular signed motifs[2]



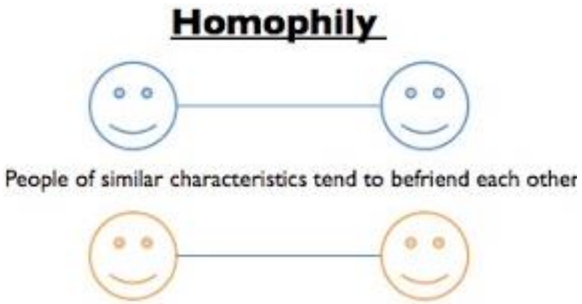


FIG.1 illustration of homophily[1]

[1] "Inequality's Economic and Social Roots: The Role of Social Networks and Homophily." Available at SSRN 3795626 (2021).
[2] "Finding, counting and listing all triangles in large graphs, an experimental study." International workshop on experimental and efficient algorithms. 2005.

Existing Signed GRL Methods

- Based on balance theory
  - SIDE, WWW 2018
  - BESIDE, CIKM 2018
  - SGCN, ICDM 2018
  - SNEA, AAAI 2020
  - ASiNE, SIGIR 2020
  - SIHG, TKDE 2020
  - SHCN, TOIS 2020
  - SGDN, Arxiv 2020
  - SGDNN, AAAI 2021

Signed GNN(using message passing mechanism)

- Not based on balance theory
  - SLF, KDD 2019
  - ROSE, WWW 2020

Most signed GRL methods are based on balance theory

Almost all signed GNN methods are based on balance theory

The Most Popular Solution: Balance Theory

- Balance theory: a well-known social theory
  - “The friend(foe) of my friend is my friend(foe)”
  - “The friend(foe) of my foe is my foe(friend)”

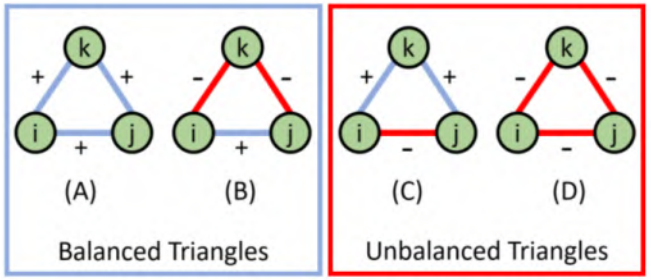


FIG.1 illustration of balance theory[1]

- Signed GRL methods based on balance theory
  - Signed Network Embedding
    - Signed random walk based on balance theory
  - Signed Graph Neural Network
    - Aggregate layer by layer based on balance theory

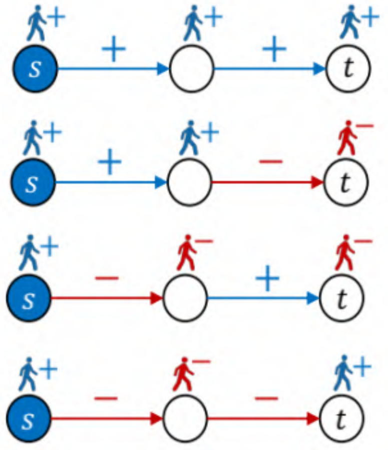


FIG.2 illustration of signed random walk[2]

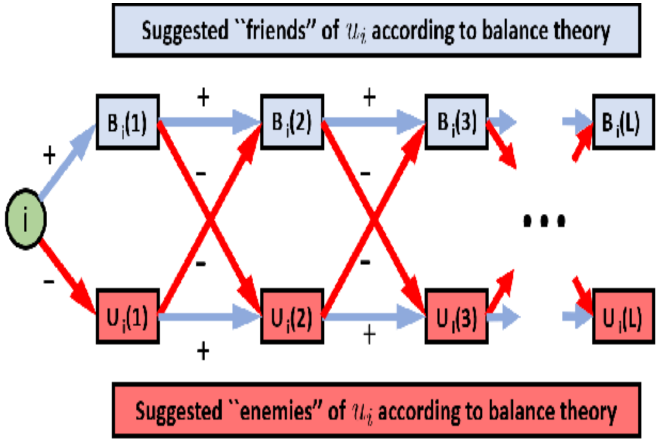


FIG.3 illustration of signed message passing[1]

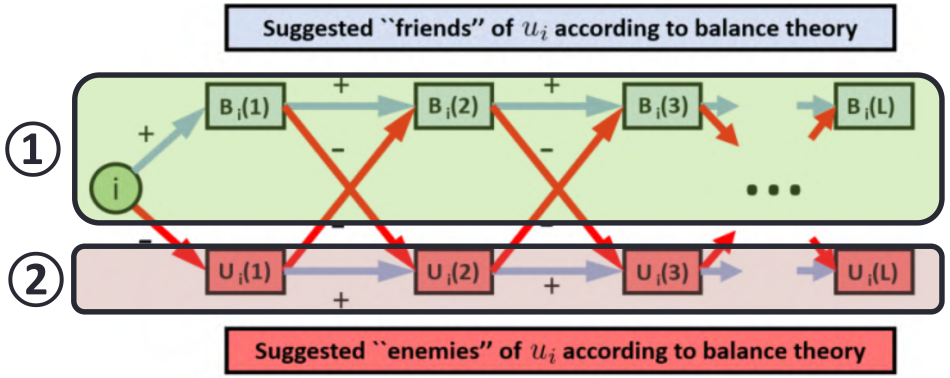
[1] Derr T, Ma Y, Tang J. Signed graph convolutional networks[C]. 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018: 929-934.
[2] Jung J, Yoo J, Kang U. Signed Graph Diffusion Network[J]. arXiv preprint arXiv:2012.14191, 2020.

Limitations of Balance Theory

- Theoretical Analysis:
  - Balance theory essentially equals to the two-conflict-groups assumption
    - Theoreom A[1]

      Let G be a signed undirected complete graph in which each triangle has an odd number of postive edges.
      Then the nodes of G can be partioned into two sets A and B (where one of A or B may be empty), such that all edges within A and B are positive, and all edges with one end in A and the other in B are negative.
    - Theoreom B[2]

      A s-graph G is balanced if and only if its point set E can be partioned into two disjoint subsets E_1, E_2, in such a way that each positve line of G joints two points of the same subset and each negative line joints two points of different subsets

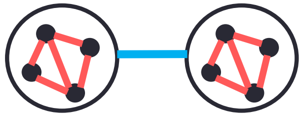


Balance theory is too ideal

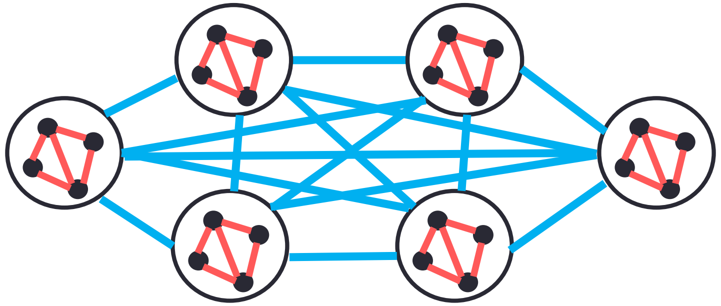
[1] Frank Harary et al. 1953. On the notion of balance of a signed graph. Michigan Mathematical Journal (1953).
[2] Cartwright D, Harary F. Structural balance: a generalization of Heider's theory[J]. Psychological review, 1956, 63(5): 277.

Beyond Balance Theory: K-group Theory

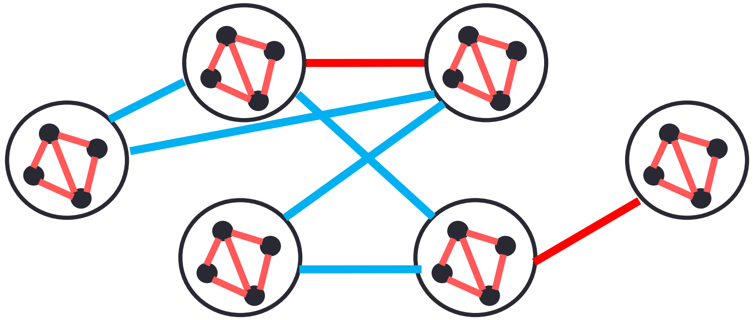
- Step 0: two conflict groups (balance theory)



- Step 1: k conflict groups

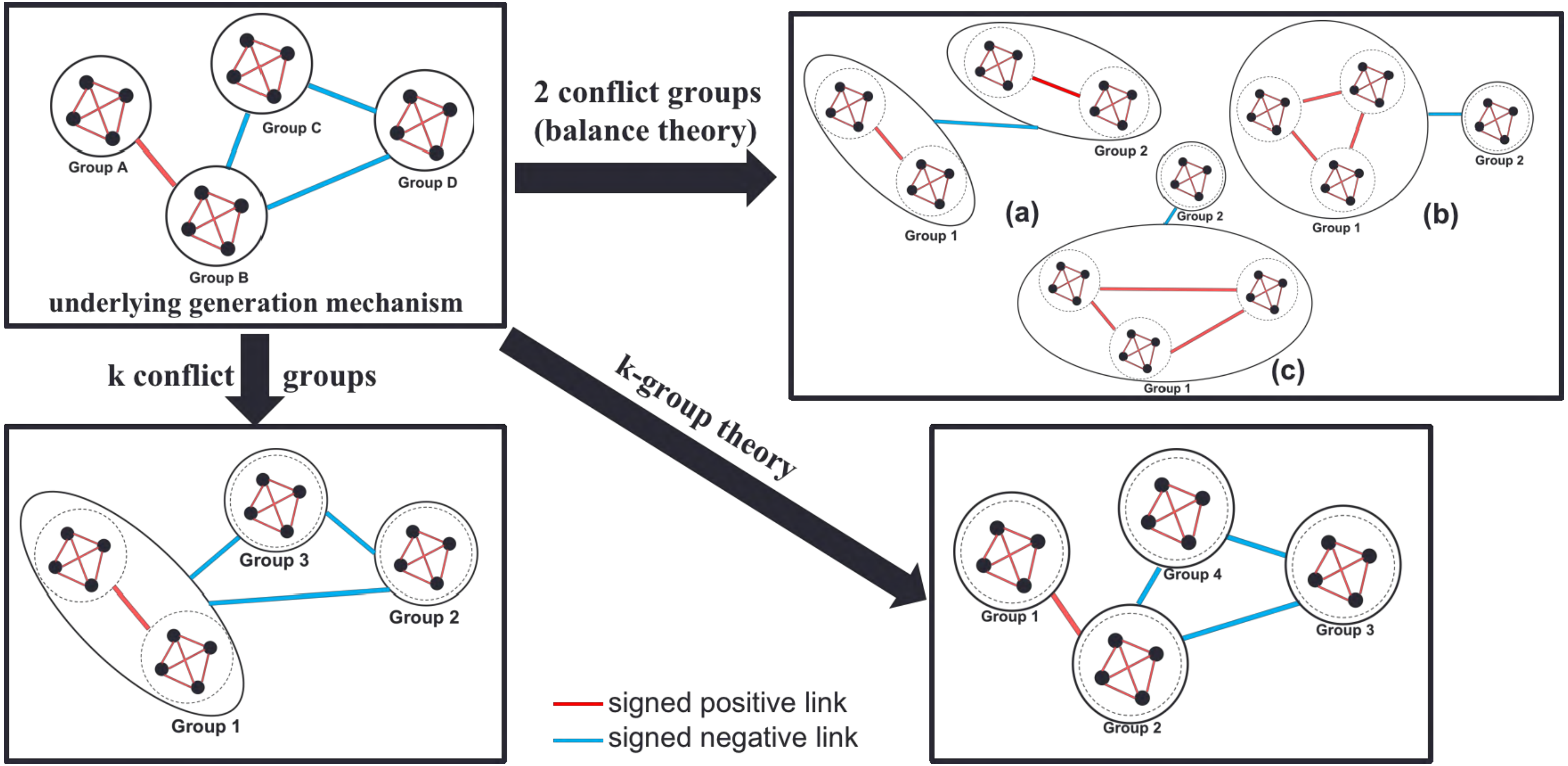


- Step2: k-group theory
  - K groups with arbitrary relations between groups
    - Negative
    - Positive
    - Netural
    - ......



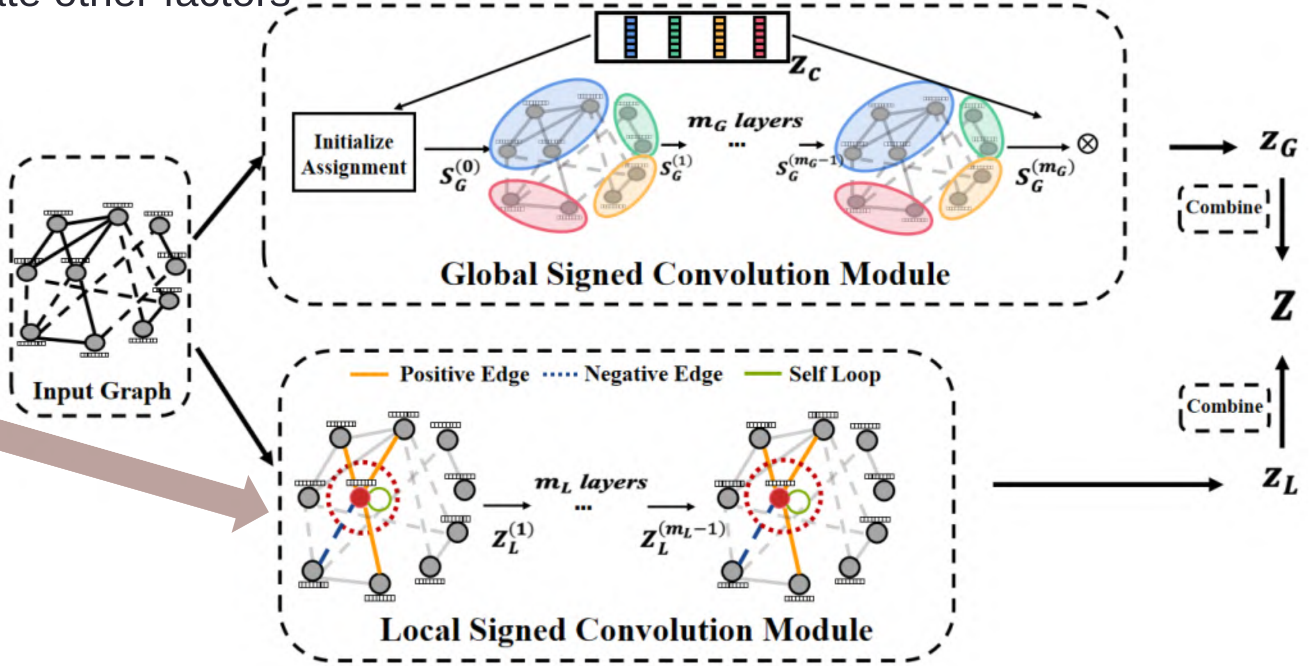
signed positive link
signed negative link

Beyond Balance Theory: K-group Theory



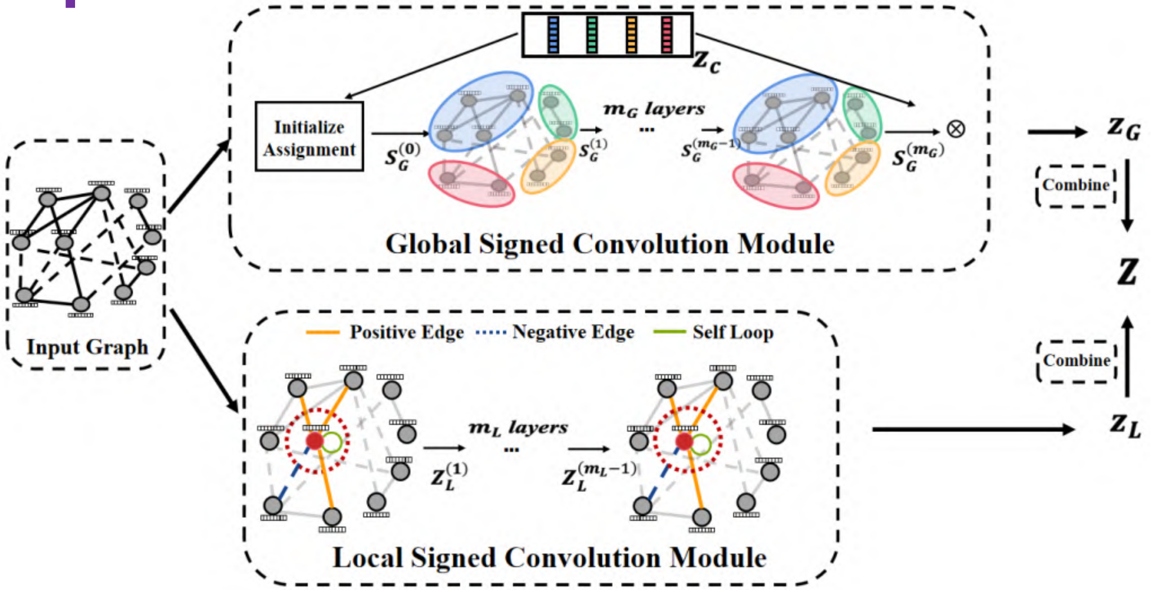
Final Assumption & Overall Framework

- Final Assumption: Combine Global and Local View
  - Global: k-group theory
    - Capture underlying k groups with arbitrary relations between groups
  - Local: without any assumption
    - Give the model more flexibility to accommodate other factors
      - Micro-structures within groups
      - Influenced by node features
      - Individual heuristic information
        e.g., always forms negative relations
      - Tolerate randommes/noise
      - ......



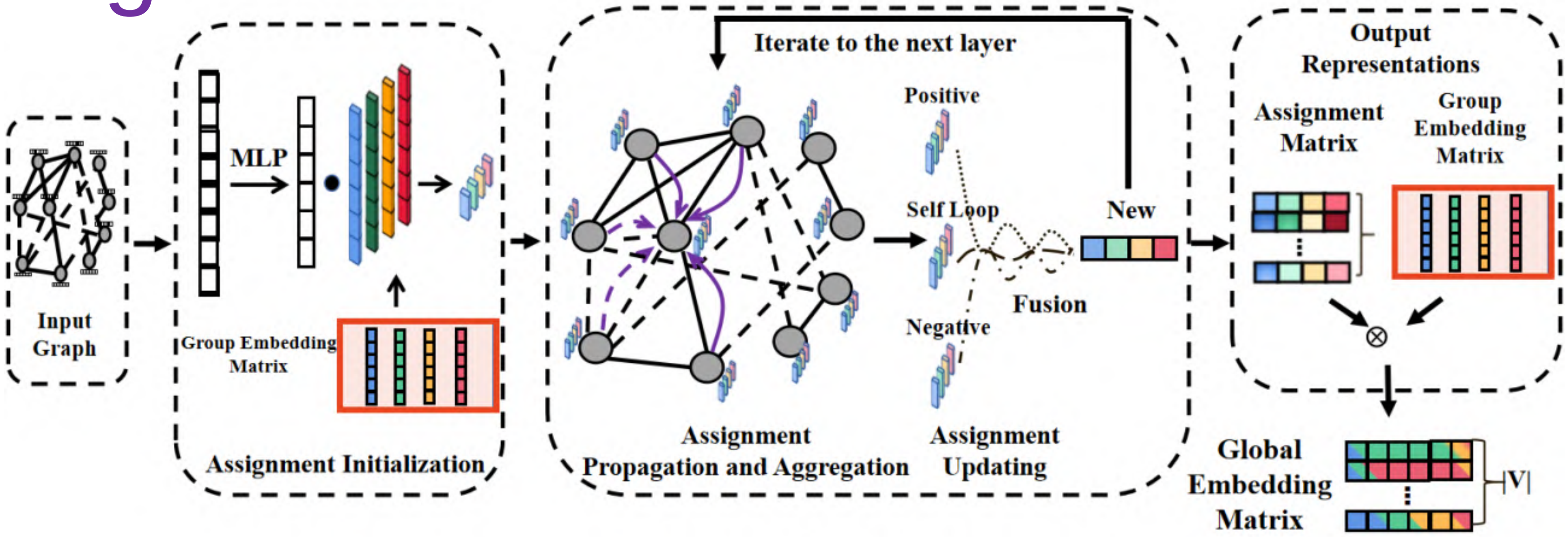
- Proposed: Group Signed GNN
  - A dual architecture
    - Global signed convolution module
    - Llobal signed convolution module

Overall Comparison



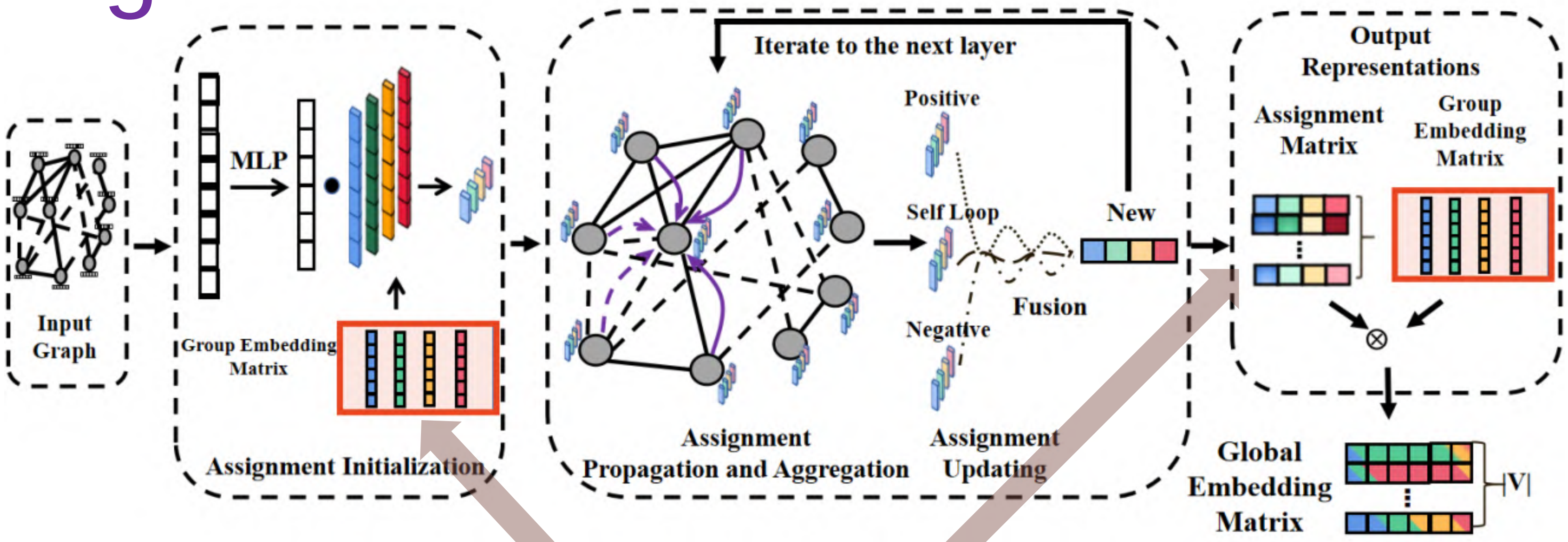
- **Assumptions**
  - 2 conflict groups(Balance Theory)
  - K conflict groups
  - **K groups with arbitrary relation**
  - **Global:K groups with arbitrary relations &Local: other flexible factors without any assumption**
- **Methods**
  - 2 conflict groups(Balance Theory)
  - K conflict groups
  - **Global model of our signed GRL model**
  - **Our signed GRL model: Group Signed GNN**

Global Signed Convolution Module



- Goal
  - Discover latent community structure based on k-group theory
- Challenges
  - Model complex relationships between communities
  - Represent nodes in a view of the groups
  - Scalable

Global Signed Convolution Module

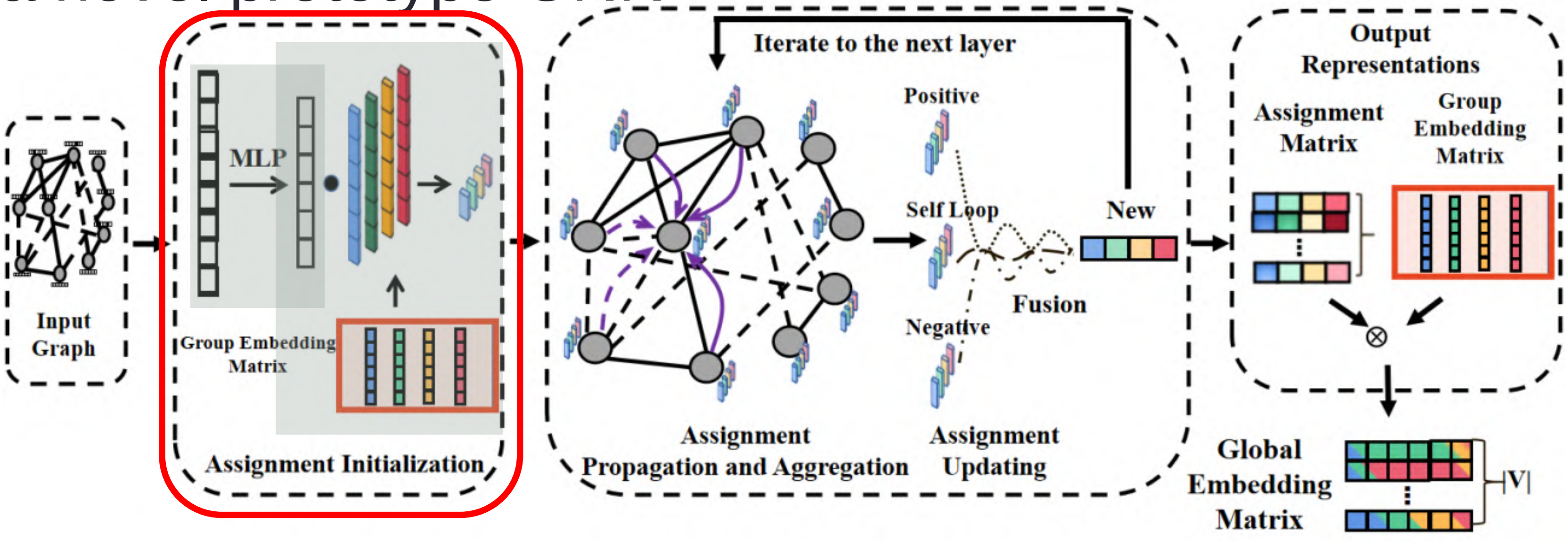


Solutions

- Denote a learnable embedding matrix $\mathbf{Z_C} = [\mathbf{Z}_{C_1}, \mathbf{Z}_{C_2}, ..., \mathbf{Z}_{C_K}] \in \mathbb{R}^{K \times d_G}$ for K groups
- Complex relationships are freely modeled in the hidden space
- Node global embeddings $\mathbf{Z}_G$ are represented as a linear combination of the group embeddings i.e.,$\mathbf{Z}_G = \mathbf{S}\mathbf{Z}_C$, where assignment matrix $\mathbf{S} \in \mathbb{R}^{N \times K}$ is learned

Global Signed Convolution Module: Details

- Model: a novel prototype GNN



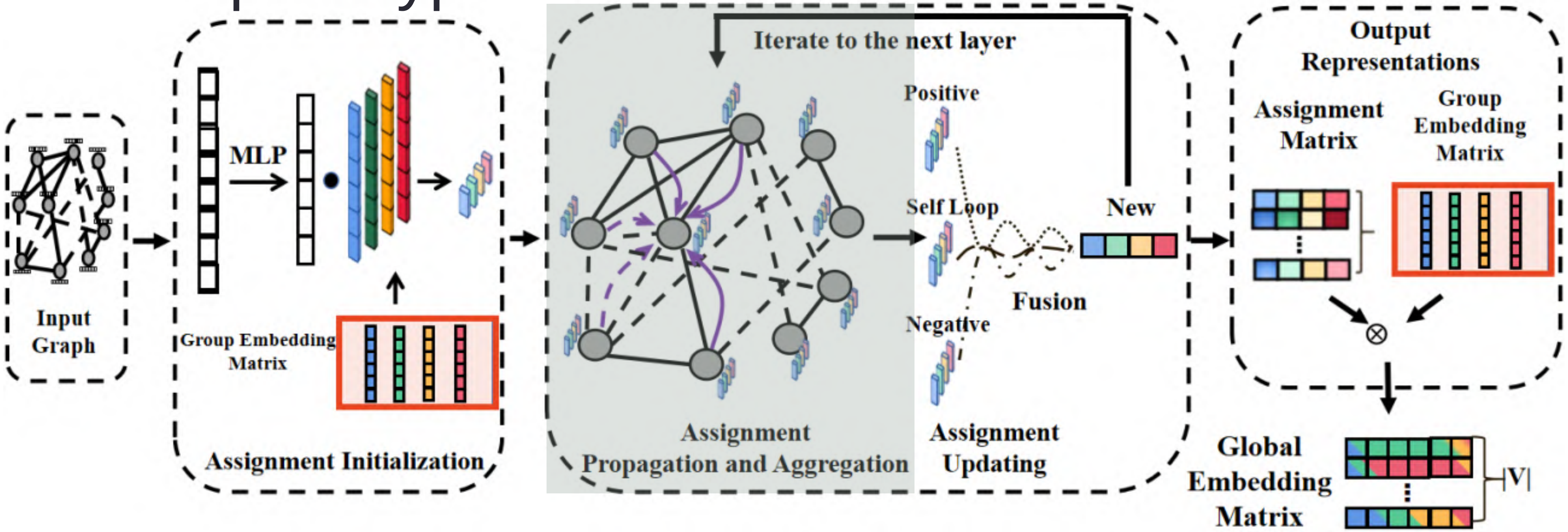
  - Assignment Initialization
    - Generate the initial assignment probability matrix $\mathbf{S} \in \mathbb{R}^{N \times K}$

$$\mathbf{X}'_v = \mathrm{MLP}(\mathbf{X}_v) \qquad (1)$$

$$\mathbf{Q}^{(0)}_{v,C_i} = \mathbf{Z}^T_{C_i}\mathbf{X}'_v, \mathbf{S}^{(0)}_{v,C_i} = \frac{\exp\left(\mathbf{Q}^{(0)}_{v,C_i}\right)}{\sum_{j=1}^{K}\exp\left(\mathbf{Q}^{(0)}_{v,C_j}\right)} \qquad (2)$$

Global Signed Convolution Module: Details

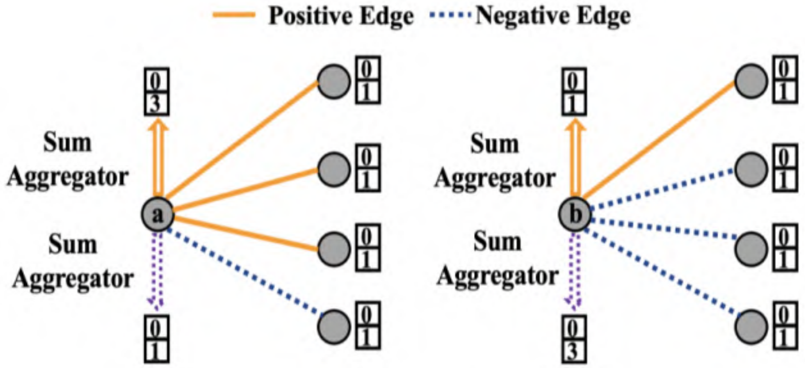
- Model: a novel prototype GNN



- Assignment Propagation and Aggregation

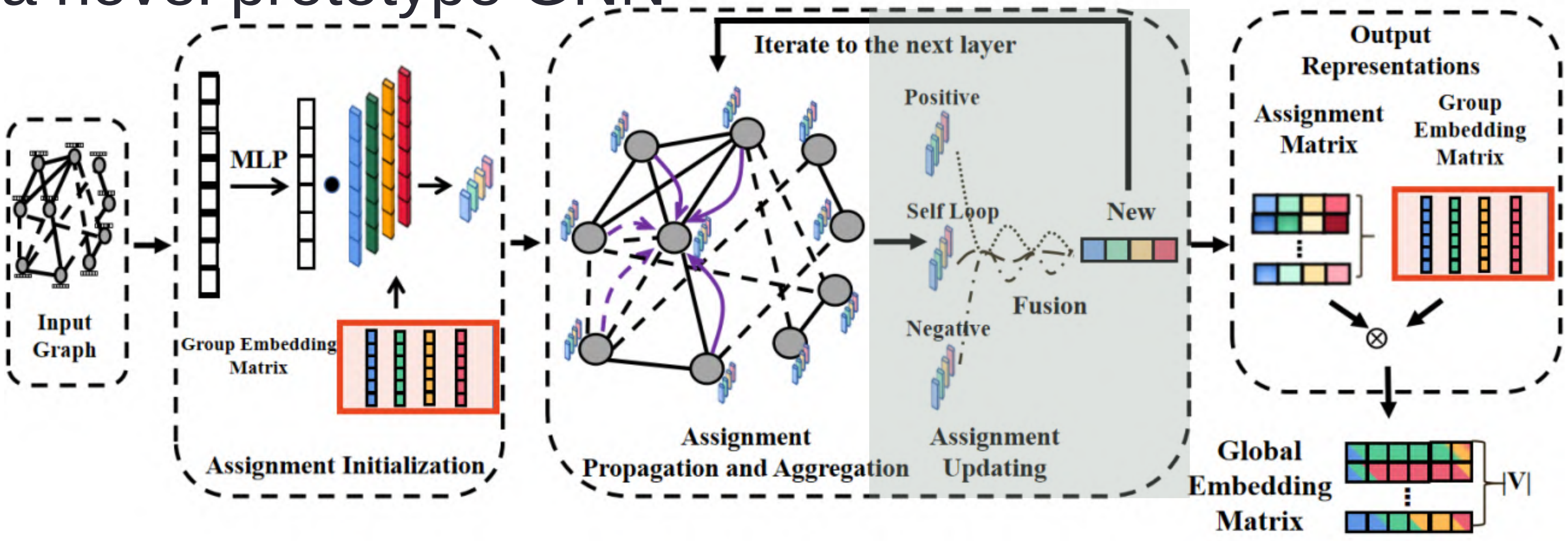
$$\mathbf{p}_v^{(m)} = \sum_{u \in \mathcal{N}_v^+} \mathbf{S}_u^{(m-1)}, \mathbf{n}_v^{(m)} = \sum_{u \in \mathcal{N}_v^-} \mathbf{S}_u^{(m-1)}$$

sum aggregator



Global Signed Convolution Module: Details

- Model: a novel prototype GNN

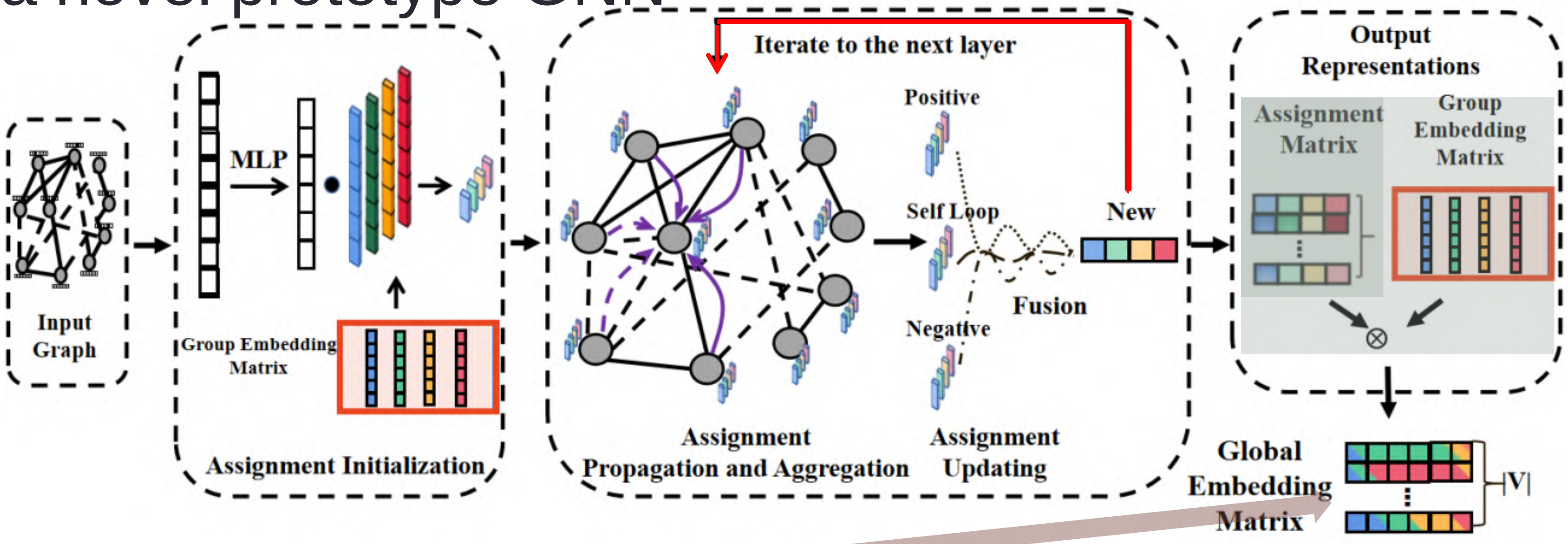


  - Assignment Updating

$$\mathbf{S}_v^{(m)} = \mathcal{F}^{(m)}\left(\mathbf{S}_v^{(m-1)}, \mathbf{p}_v^{(m)}, \mathbf{n}_v^{(m)}\right)$$
$$= \mathrm{softmax}\left(\sigma\left(\left[\mathbf{S}_v^{(m-1)}, \mathbf{p}_v^{(m)}, \mathbf{n}_v^{(m)}\right]\mathbf{W}_G^{(m)'}\right)\mathbf{W}_G^{(m)}\right)$$

Global Signed Convolution Module: Details

- Model: a novel prototype GNN



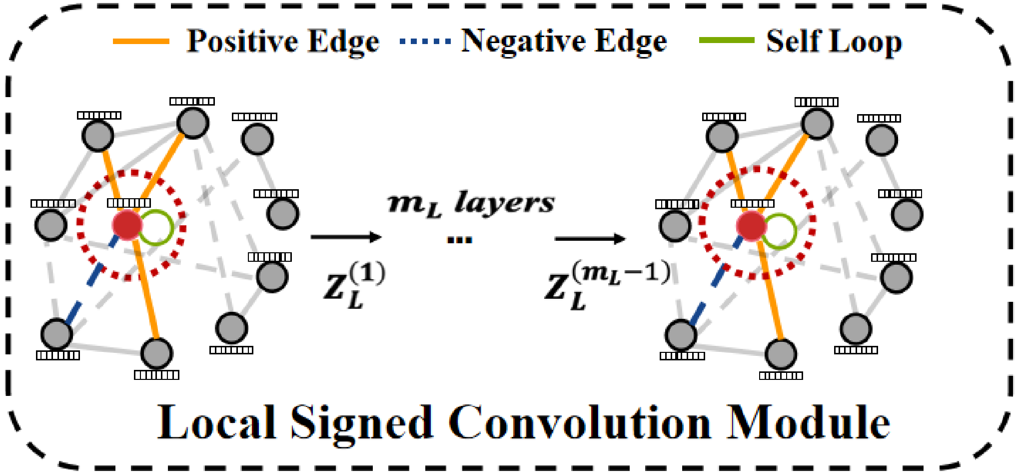
- Obtaining Global Representations $\mathbf{Z}_G$
  - Repeating the 2nd ans 3rd steps for M_G time, i.e., adopting M_G message-passing layers
  - Obtaining the final assignment matrix $\mathbf{S}$

    $$\mathbf{S} = \mathbf{S}^{(M_G)}$$

  - Using the linear combination of group embeddings

    $$\mathbf{Z}_G = \mathbf{S}\mathbf{Z}_C$$

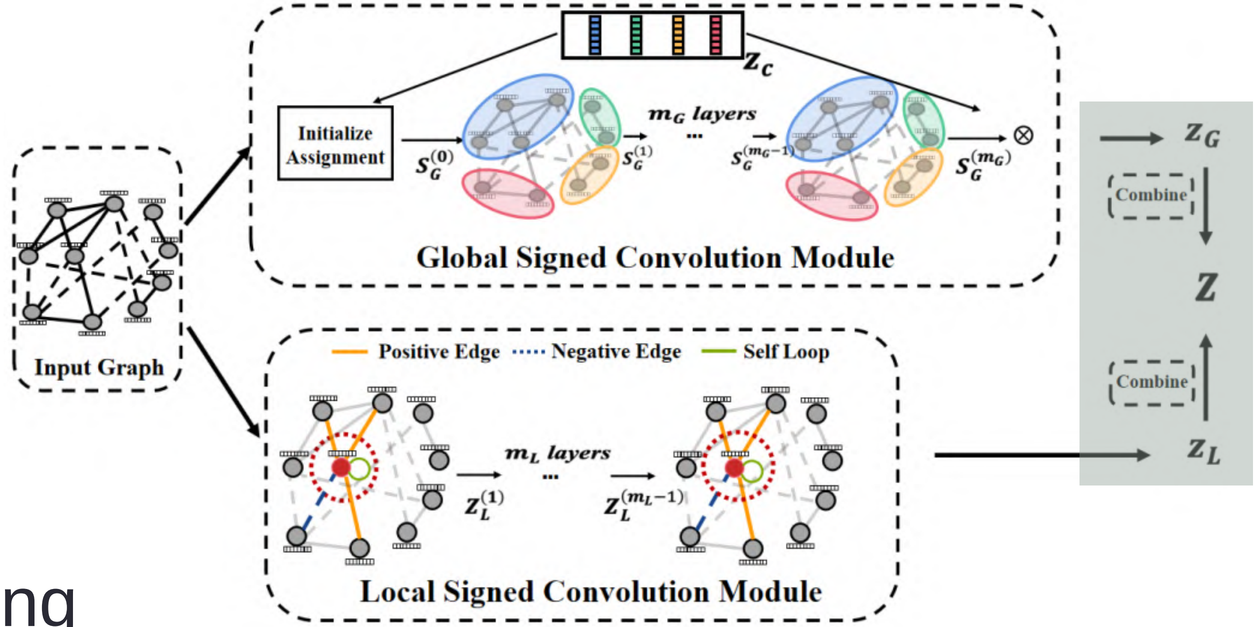
Local Signed Convolution Module



- Goal
  - Give the model more flexibility to accommodate other factors
  - Without any assumption
- Solution
  - Treat self connections, positive links, and negative links as three relations
- Model
  - A multi-relational GNN
  - Conduct message-passing for M_L layers, the final node local embeddings are $\mathbf{Z}_L$

$$\mathbf{h}_v^{(m)} = \left[\sum\nolimits_{u \in \mathcal{N}_v^+} \mathbf{z}_u^{(m-1)}, \sum\nolimits_{u \in \mathcal{N}_v^-} \mathbf{z}_u^{(m-1)}\right], \quad \mathbf{z}_v^{(m)} = \sigma\left(\left[\mathbf{z}_v^{(m-1)}, \mathbf{h}_v^{(m)}\right] \mathbf{W}_L^{(m)} + \mathbf{b}_L^{(m)}\right)$$

Proposed Model: Group Signed GNN (GS-GNN)



- Final Embedding
  - Concat global embeddings $\mathbf{Z}_G$ and local embeddings $\mathbf{Z}_L$
    $\mathbf{Z} = [\mathbf{Z}_G, \mathbf{Z}_L]$
- Objective function
  - e.g. for link sign prediction task, i.e., predicting the polar of the given links

$$\mathcal{L} = -\frac{1}{|\mathcal{E}^+ \cup \mathcal{E}^-|}\left(\sum_{(u,v)\in\mathcal{E}^+} \log p(u,v) + \sum_{(u,v)\in\mathcal{E}^-} (1 - \log p(u,v))\right) + \lambda\mathcal{L}_{\text{reg}}$$

Experimental Setting: Datasets, Task & Metrics

- Datasets
  - Four public real-world signed graphs
    - Bitcoin-Alpha, Bitcoin-OTC: two signed graphs extracted from bitcoin trading platforms
    - Slashdot: a technology-related news website
    - Epinions: a consumer review site
  - Synthetic datasets
- Task
  - Link sign prediction
- Evaluation Metrics
  - Four metrics
    - AUC: area under curve
    - Macro-F1:macro-averaged F1 score
    - Micro-F1: micro-average F1 score
    - Binary-F1: binary average F1 score
  - Higher value indicates better performance

Table 1: The Statistics of Real-world Datasets

Datasets	# Nodes	# Links	# Positive Links (Ratios)	# Negative Links (Ratios)
Bitcoin-Alpha	3,775	14,120	12,721 (90.09%)	1,399 (9.91%)
Bitcoin-OTC	5,875	21,489	18,230 (84.83%)	3,259 (15.17%)
Slashdot	37,626	419,072	313,543 (74.82%)	105,529 (25.14%)
Epinions	45,003	616,031	513,851 (83.41%)	102,180 (16.59%)

Experimental Setting: Baselines

- Baselines
  - Signed graph clustering method
    - SPONGE (AISTATS 2019)
  - Signed network embedding based on balance theory
    - SIDE (WWW 2018)
  - Signed network embedding not based on balance theory
    - SLF (KDD 2019)
  - Unsigned GNN
    - GCN (ICLR 2017)
  - Signed GNN based on balance theory
    - SGCN (ICDM 2018)
    - SNEA (AAAI 2020): +attention
    - SGDN (Arxiv 2020): +diffusion
- Our method: GS-GNN

Results on Synthetic Dataset

- Question 1
  - Can GS-GNN fully utilize the k-group theory and discover the underlying structure of signed graphs?
- Synthetic dataset
  - Using the signed stochastic block model (SSBM) to generate K_S conflict groups with random noise
- Results of Macro-F1

Assumption		Method	$K_S = 2$	$K_S = 3$	$K_S = 4$	$K_S = 5$	$K_S = 6$
Balance Theory		SGCN	0.442	0.398	0.362	0.334	0.357
		SGDN	0.791	0.682	0.612	0.530	0.495
K-Group	K=K_S	SPONGE	0.983	0.989	0.990	0.990	0.989
		GS-GNN	0.984	0.991	0.991	0.989	0.982
	K=2	SPONGE	0.983	0.853	0.749	0.670	0.600
		GS-GNN	0.984	0.991	0.988	0.984	0.889
	K=6	SPONGE	0.463	0.662	0.848	0.940	0.989
		GS-GNN	0.986	0.988	0.990	0.989	0.980

- Conclusion
  - Demonstrate the superiority of GS-GNN in utilizing the k-group theory
  - GS-GNN even outperforms the SPONGE

Results on Real Graphs

- Question 2
  - How does GS-GNN perform on different real graphs which is usually complicated, compared with other state-of-the-art signed graphs representation learning methods?

- Results

Dataset	Metric	SPONGE	SLF	SIDE	GCN	SGCN	SNEA	SGDN	GS-GNN	
Bitcoin-Alpha	AUC	0.513	0.847	0.797	0.806	0.858	0.866	0.840	**0.893**	+3.12%
	Macro-F1	0.504	0.668	0.665	0.546	0.706	0.727	0.663	**0.793**	+9.08%
	Micro-F1	0.901	0.819	0.824	0.902	0.864	0.873	0.894	**0.930**	+3.10%
	Binary-F1	0.948	0.892	0.896	0.948	0.921	0.926	0.942	**0.961**	+1.37%
Bitcoin-OTC	AUC	0.700	0.873	0.828	0.845	0.871	0.863	0.863	**0.915**	+4.81%
	Macro-F1	0.644	0.735	0.713	0.675	0.754	0.760	0.734	**0.837**	+10.13%
	Micro-F1	0.763	0.828	0.820	0.875	0.850	0.858	0.871	**0.920**	+5.14%
	Binary-F1	0.850	0.892	0.889	0.928	0.908	0.914	0.926	**0.952**	+2.59%
Slashdot	AUC	0.500	0.888	0.820	0.819	0.873	0.888	0.887	**0.916**	+3.15%
	Macro-F1	0.432	0.772	0.725	0.670	0.760	0.769	0.769	**0.812**	+5.18%
	Micro-F1	0.752	0.812	0.773	0.797	0.802	0.812	0.838	**0.865**	+3.22%
	Binary-F1	0.861	0.867	0.840	0.875	0.859	0.868	0.896	**0.915**	+2.12%
Epinions	AUC	0.508	0.928	0.878	0.869	0.925	0.931	0.930	**0.959**	+3.01%
	Macro-F1	0.474	0.795	0.746	0.685	0.800	0.819	0.819	**0.865**	+5.62%
	Micro-F1	0.832	0.865	0.829	0.864	0.872	0.888	0.903	**0.931**	+3.10%
	Binary-F1	0.908	0.915	0.891	0.922	0.920	0.931	0.942	**0.961**	+2.02%

- Observations
  - SPONGE fails on real-world graphs
  - Unsigned GNN outperforms other baselines in some cases
  - GS-GNN consistently outperforms all the baselines on all datasets with all evaluation metrics
  - 5.14%~10.13% improvement of Macro-F1 indicates that GS-GNN can better model the negative links

Ablation Study Results

- Question 3
  - Does sum aggregator contribute to our proposed GS-GNN method?
- Results
  - The ablation study results of sum aggregator

Dataset	Metric	$\text{GS-GNN}_{\text{mean}}$	$\text{GS-GNN}_{\text{sum}}$
Bitcoin-Alpha	AUC	0.844	0.893
	Macro-F1	0.712	0.793
	Micro-F1	0.915	0.930
	Binary-F1	0.954	0.961
Bitcoin-OTC	AUC	0.900	0.915
	Macro-F1	0.812	0.837
	Micro-F1	0.914	0.920
	Binary-F1	0.950	0.952

- Conclusion
  - Using the sum aggregator for positive and negative neighbors separately in signed graphs is important

Ablation Study Results

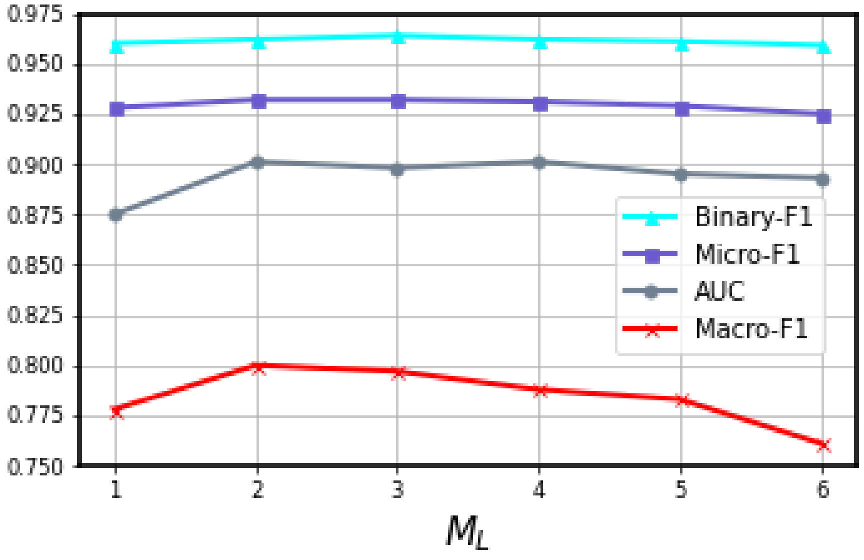
- Question 3
  - Do both the global and local representation contribute to our proposed GS-GNN method?
- Results
  - The ablation study results of local and global representation

Dataset	Metric	GS-GNN$_L$	GS-GNN$_G$	GS-GNN
Bitcoin-Alpha	AUC	0.875	0.889	0.893
	Macro-F1	0.754	0.731	0.793
	Micro-F1	0.922	0.916	0.930
	Binary-F1	0.958	0.954	0.961
Bitcoin-OTC	AUC	0.891	0.906	0.915
	Macro-F1	0.801	0.786	0.837
	Micro-F1	0.901	0.899	0.920
	Binary-F1	0.946	0.942	0.952

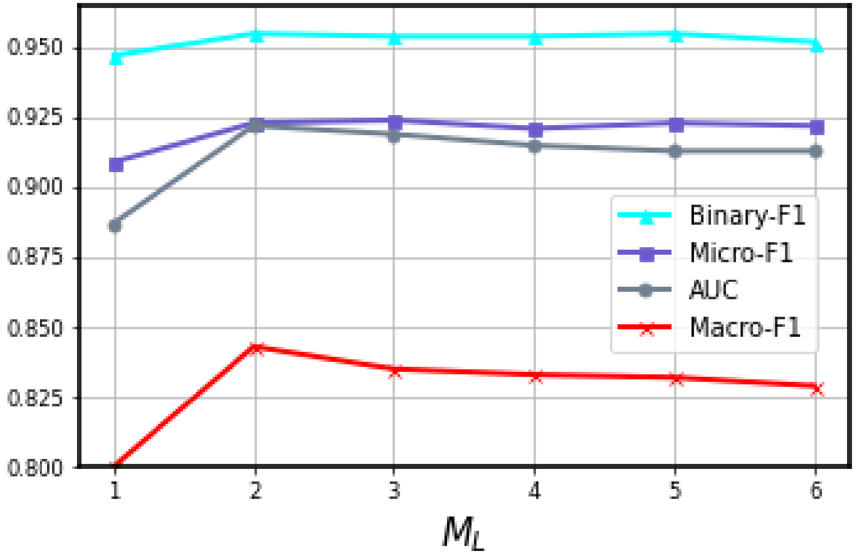
- Conclusion
  - Both modules contribute to GS-GNN
  - The local and global representations of nodes are complementary

Parameter Sensitivities Results

- Question 4
  - How do essential parameters affect the model?
- Results
  - Varying the number of layers $\mathbf{M}_L$ in the local module



Bitcoin-Alpha

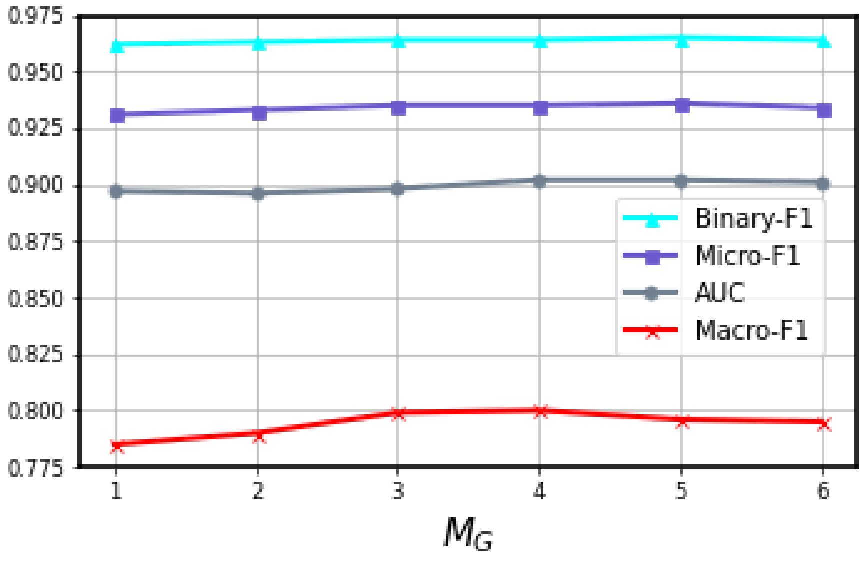


Bitcoin-OTC

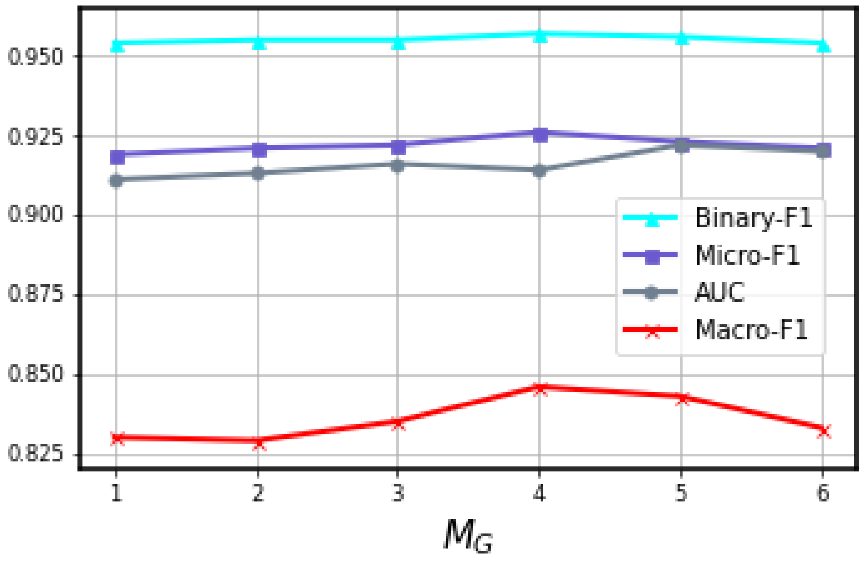
- Conclusion
  - 2 local layers is a suitable choice

Parameter Sensitivities Results

- Question 4
  - How do essential parameters affect the model?
- Results
  - Varying the number of layers $\mathbf{M}_G$ in the golal module



Bitcoin-Alpha

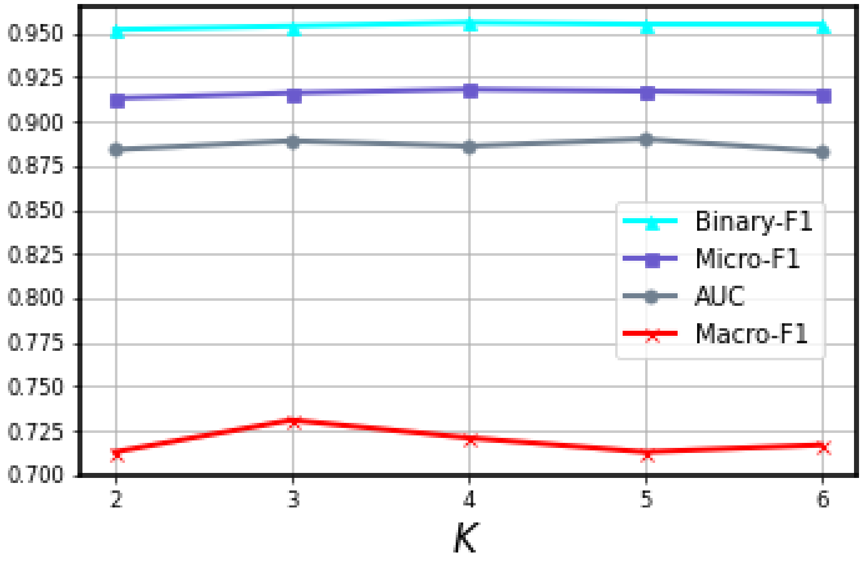


Bitcoin-OTC

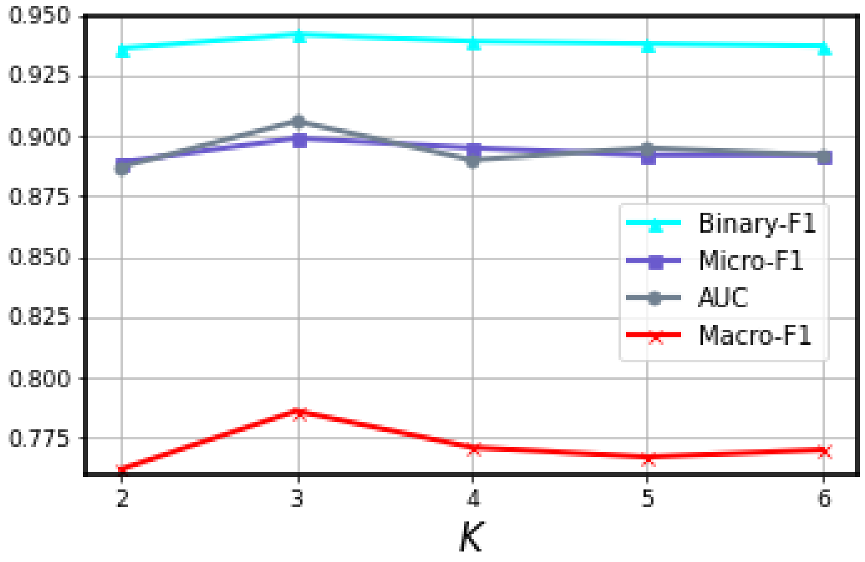
- Conclusion
  - 5 gobal layers is a suitable choice

Parameter Sensitivities Results

- Question 4
  - How do essential parameters affect the model?
- Results
  - Varying the number of groups K



Bitcoin-Alpha



Bitcoin-OTC

- Conclusion
  - Setting K from 3 to 5 leads to the best results

Conclusion

- Study representation learning methods for signed graphs
  - Most existing methods are based on balance theory, ignore its serious limitation
  - We propose the k-group theory
    - a general and more realistic assumption beyond the usual balance theory
- Propose a novel signed GNN with a dual architecture (GS-GNN)
  - Simultaneously learn global and local representations.
    - fully leverage the k-group theory
    - with the flexibility to capture extra information beyond k-group theory
  - Simple and effective
- Extensive experimental results on synthetic and real signed graphs
  - Demonstrate the superiority of our proposed assumption and method
  - Achieves new state-of-the-art, to the best of our knowledge

Thanks!

Haixin Liu, Tsinghua University
liu-hx20@mails.tsinghua.edu.cn