# Signed Graph Attention Networks

Junjie Huang[1,2,*], Huawei Shen[1], Liang Hou[1,2], and Xueqi Cheng[1]

[1] CAS Key Laboratory of Network Data Science and Technology,
Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
{huangjunjie17s, shenhuawei, houliang17z, cxq}@ict.ac.cn

**Abstract.** Graph or network data is ubiquitous in the real world, including social networks, information networks, traffic networks, biological networks and various technical networks. The non-Euclidean nature of graph data poses the challenge for modeling and analyzing graph data. Recently, Graph Neural Network (GNNs) are proposed as a general and powerful framework to handle tasks on graph data, e.g., node embedding, link prediction and node classification. As a representative implementation of GNNs, Graph Attention Networks (GATs) are successfully applied in a variety of tasks on real datasets. However, GAT is designed to networks with only positive links and fails to handle signed networks which contain both positive and negative links. In this paper, we propose Signed Graph Attention Networks (SiGATs), generalizing GAT to signed networks. SiGAT incorporates graph motifs into GAT to capture two well-known theories in signed network research, i.e., balance theory and status theory. In SiGAT, motifs offer us the flexible structural pattern to aggregate and propagate messages on the signed network to generate node embeddings. We evaluate the proposed SiGAT method by applying it to the signed link prediction task. Experimental results on three real datasets demonstrate that SiGAT outperforms feature-based method, network embedding method and state-of-the-art GNN-based methods like signed graph convolutional network (SGCN).

**Keywords:** Signed Network · Network Embedding · Graph Neural Network

## 1 Introduction

Graph Neural Networks (GNNs) have been receiving more and more research attentions, achieving good results in many machine learning tasks e.g., semi-supervised node classification [11], network embedding [12] and link prediction. GNNs introduce neural networks into graph data by defining convolution [11], attention [23] and other mechanisms. However, previous GNNs methods mainly focus on undirected and unsigned networks (i.e., networks consisting of only positive edges). How to apply GNNs to signed, directed, weighted and other complex networks is an important research direction.

With the growing popularity of online social media, many interactions are generated by people on the Web. Most of these interactions are positive relationships, such as friendship, following, support and approval. Meanwhile, there are also some negative links that indicate disapproval, disagreement or distrust. Some recent researches about signed social networks suggest that negative links bring more valuable information over positive links in various tasks [14]. In the early years, the methods for modeling signed networks were mainly based on feature extraction [14], matrix factorization [8] and network embedding [10,20,24]. They have achieved good performances in tasks such as signed link prediction and node classification. The core of modeling signed networks is two important sociological theories, i.e., balance theory and status theory. Derr et al. [2] proposed a signed graph convolution network (SGCN) method to model signed networks. It utilizes balance theory to aggregate and propagate the information. However, it only applies to the undirected signed networks.

In this paper, we try to model the directed signed network into graph attention network (GAT). To the best of our knowledge, it's the first time to introduce GAT to directed signed networks. Attention mechanism can characterize the different effects of different nodes on the target node, e.g., A is is more kind to me than B.

The main contributions of this paper are as follows:

- We analyze the key elements of the signed link prediction problem. Triads motifs are used to describe two important sociological theories, i.e., balance theory and status theory.
- We introduce the GAT to model the signed network and design a new motif-based GNN model for signed networks named SiGAT.
- We conduct experiments on some real signed social network data sets to demonstrate the effectiveness of our proposed framework SiGAT.

The rest of paper is organized as follows. In section 2, we review some related work. We propose our model SiGAT in section 3, which consists of introducing related social theory, incorporating these theories into motifs and finally presenting our Signed Graph Attention Network (SiGAT). In section 4, experiments are performed to empirically evaluate the effectiveness of our framework for learning node embedding. We discuss the results of some signed network embedding and SGCN, analyze the impact of hyper-parameters in our model. Finally, we conclude and discuss future work in section 5.

## 2 Related work

### 2.1 Signed Network Embedding

Signed social networks are such social networks in which social ties can have two signs: positive and negative [15]. To mine signed networks, many algorithms have been developed for tasks such as community detection, node classification, link prediction and spectral graph analysis. In particular, [15] proposed a feature engineering-based methods to do the signed link prediction tasks. Recently, with the development of network representation learning [5,17,21], researchers begin to learn low-dimensional vector representations for a social network. For signed network embedding methods, SNE [26] adopts the log-bilinear model and incorporates two signed-type vectors to capture the positive or negative relationship of each edge along the path. SiNE [24] designs a new objective function guided by social theories to learn signed network embedding, it proposed to add a virtual node to enhance the training process. SIDE [10] provides a linearly scalable method that leverages balance theory along with random walks to obtain the low-dimensional vector for the directed signed network. SIGNet [9] combines balance theory with a specialized random and new sampling techniques in directed signed networks. These methods are devoted to defining an objective function that incorporates sociological theory and then using some machine learning techniques e.g., sampling and random walks to optimize look-up embedding.

### 2.2 Graph Neural Networks

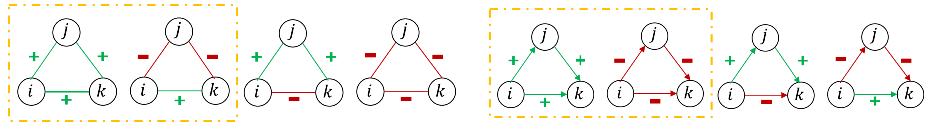
GNNs have achieved a lot of success in the semi-supervised node classification task [11], graph classification [4] and other graph analysis problems. The concept of graph neural network (GNN) was first proposed in [18], which extended neural networks for processing the data represented in graph domains. Recently, researchers tried to apply convolution [11], attention [23] , auto-encoder [12] and other mechanisms into graphs.

GCN in [11] is designed with the focus to learning representation at the node level in the semi-supervised node classification task. GraphSAGE [6] extends it to the large inductive graphs. GAE [12] firstly integrates GCN into a graph auto-encoder framework. This approach is aimed at representing network vertices into a low-dimensional vector space by using neural network architectures in unsupervised learning tasks. GAT [23] applies attention mechanisms to GNNs. Compared to the convolution, the attention mechanism can describe the influence of different nodes on the target node. On the semi-supervised task of the node, it shows better results than GCN. SGCN [1] introduced GNNs to a signed network for the first time. It designed a new information aggregation and propagation mechanism for the undirected signed network according to the balance theory. However, the majority of current GNNs tackle with static homogeneous graphs [25], how to apply GNNs into other complex networks(e.g., directed signed networks) is still a challenge.

## 3 Methods

### 3.1 Signed Network Theory

Balance theory and status theory play an important role in analyzing and modeling signed networks. These theories can enable us to characterize the differences between the observed and predicted configurations of positive and negative links in online social networks [14].



(a) Illustration of structural balance theory

(b) Illustration of status theory

**Fig. 1.** The balance theory and status theory

**Balance Theory** Balance theory originated in social psychology in the mid-20th-century [7]. It was initially intended as a model for undirected signed networks. All signed triads with an even number of negative edges are defined as balanced. In other words, for the four triangles in Fig.1(a), the triangles which all three of these users are friends or only one pair of them are friends are defined as balanced i.e., the first two triads are balanced. Balance theory suggests that people in a social network tend to form into a balanced network structure. It can be described as "the friend of my friend is my friend" and "the enemy of my enemy is my friend". This theory has a wide range of applications in the field of sociology. [14] found that balance theory has widely existed in social networks. It was proven that if a connected network is balanced, it can be split into two opposing groups and vice versa [3].

**Status Theory** Status theory is another key social psychological theory for signed network. It is based on directed signed networks. It supposes directed relationship labeled by a positive sign "+" or negative sign "-" means target node has a higher or lower status than source node [22]. For example, a positive link from A to B, it does not only mean "B is my friend" but also "I think B has a higher status than I do.". For the triangles in Fig.1(b), the first two triangles satisfy the status ordering and the latter two do not satisfy it. For the first triangle, when Status(j) > Status(i) and Status(k) > Status(j), we have Status(k) > Status(i). Based on status theory, it can be found that "opinion leaders" have higher social status (manager or advisor) than ordinary users [22].

Based on the analysis of the above two theories on social network datasets, [14] extracts a total of 23 features of two given nodes $u$ and $v$ and then uses logistic regression to train and infer the positive and negative of edges. Experiments show that these features can help to achieve good results in machine learning tasks. These two theories can be described in triangles. These directed edges, signed edges and triangles can all be defined as motifs of a network [16]. In signed social networks, nodes not only spread the message/information/influence by just directed positive or negative relationship but also triad motifs.

### 3.2 Signed Motifs

Since positive neighbors and negative neighbors should have different effects on the target node, they should obviously be treated separately. Furthermore, for directed signed networks, the direction also contains some knowledge, e.g., status theory. Regard as previous discussions about social-psychological theories, we think that triads also have different effects. As we said, these different effects can be described by different motifs. Therefore, we propose a unified framework to effectively model signed directed networks based on motifs.

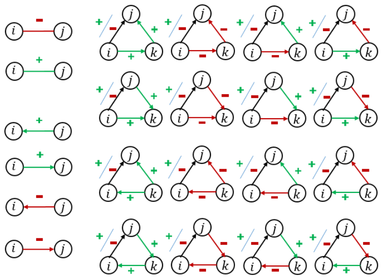


**Fig. 2.** 38 different motifs in SiGAT, mean different influences from node $j$ to node $i$

In our framework, we define positive/negative neighbors (2 motifs), positive /negative with direction neighbors (4 motifs) and 32 different triangle motifs in Fig.2. We first extract these motifs and then apply them to our model. The method of effectively extracting these motifs can be found in [19].

### 3.3 Signed Graph Attention Networks

GAT [23] introduces attention mechanism into the graph. It uses a weight matrix to characterize the different effects of different nodes on the target node. GAT firstly computes the $\alpha_{ij}$ for node $i$ and node $j$ by the attention mechanism $\mathbf{a}$ and LeakyReLU nonlinearity (with negative input slope $\alpha = 0.2$) as:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\mathbf{a}^T[\mathbf{W}\boldsymbol{h}_i \| \mathbf{W}\boldsymbol{h}_j]\right)\right)}{\sum_{k \in \mathcal{N}_i} \exp\left(\text{LeakyReLU}\left(\mathbf{a}^T[\mathbf{W}\boldsymbol{h}_i \| \mathbf{W}\boldsymbol{h}_k]\right)\right)}, \tag{1}$$

where $\cdot^T$ represents transposition and $\|$ is the concatenation operation, $\mathcal{N}_i$ are the neighborhoods of node $i$, $\mathbf{W}$ is the weight matrix parameter, $\boldsymbol{h}_i$ is the node feature of node $i$.

After computing the normalized attention coefficients, a linear combination of the features is used and served to the final output features for every node:

$$\boldsymbol{h}'_i = \sigma\left(\sum_{j \in \mathcal{N}_i} \alpha_{ij} \mathbf{W}\boldsymbol{h}_j\right). \tag{2}$$

Based on previous discussions, we proposed our SiGAT model in Fig.3. For a node $u$, in different motifs, we get different neighborhoods. For example, in $v \rightarrow^{+} u$ motifs definition, the gray color nodes $v$ are the neighborhoods, we use

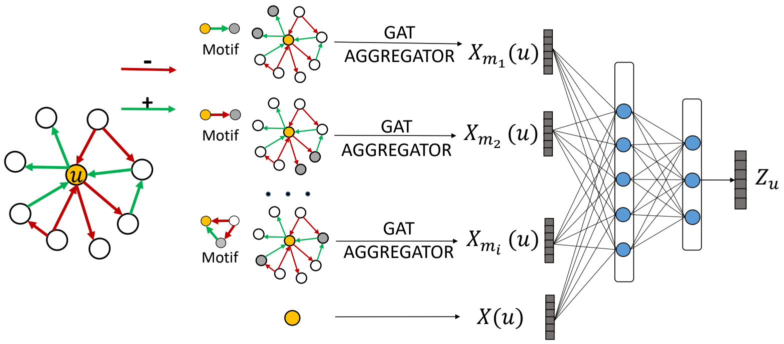


**Fig. 3.** For a signed network, the target $u$ is the orange node. Red/Green links mean positive/negative links, the arrows indicate the directions. We define different motifs and obtain the corresponding neighborhoods in them. Under a motif definition, the gray color nodes are the neighborhoods, we use GAT to describe the influences from these neighborhoods. After aggregating the information from different motifs, we use these $\mathbf{X}_{m_i}(u)$ and $\mathbf{X}(u)$ to get the representation $\mathbf{Z}_u$.

GAT to describe the influence from these neighborhoods. After aggregating the information from different motifs, we concatenate these $\mathbf{X}_{m_i}(u)$ and $\mathbf{X}(u)$ to a two-layer fully connected neural network to generate the representation $\mathbf{Z}_u$ for node $u$, $m_i$ is the $i$-th motif. We also take the scalability into considerations and propose a mini-batch algorithm like GraphSAGE [6].

The details are described in Algorithm 1. First, for each node $u$, we initialize embeddings $\mathbf{X}(u)$ as features. We extracted new graphs under different motif definitions using the motif graph extract function $F_{m_i}$, based on motif list which is given before. In other word, for each target node $u$, we get its neighbor nodes $\mathcal{N}_{m_i}(u)$ under motif $m_i$ definition. Then, for every motif $m_i$, we use a GAT model GAT-AGGREGATOR$_{m_i}$ with the parameter $\mathbf{W}_{m_i}, \boldsymbol{a}_{m_i}$ as follows to get the message $\mathbf{X}_{m_i}(u)$ in this motif:

$$\alpha_{uv}^{m_i} = \frac{\exp\left(\text{LeakyReLU}\left(\mathbf{a}_{m_i}^T[\mathbf{W}_{m_i}\mathbf{X}(u)\|\mathbf{W}_{m_i}\mathbf{X}(v)]\right)\right)}{\sum_{k\in\mathcal{N}_{m_i}(u)}\exp\left(\text{LeakyReLU}\left(\mathbf{a}_{m_i}^T[\mathbf{W}_{m_i}\mathbf{X}(u)\|\mathbf{W}_{m_i}\mathbf{X}(k)]\right)\right)}, \tag{3}$$

$$\mathbf{X}_{m_i}(u) = \sum_{v\in\mathcal{N}_{m_i}(u)} \alpha_{uv}^{m_i}\mathbf{W}_{m_i}\mathbf{X}(v). \tag{4}$$

Finally, we concatenate all the messages from different motifs with $\mathbf{X}(u)$ to a two-layer fully connected neural network to obtain the final embedding $\mathbf{Z}_u$ for the target node $u$. Here, $\mathbf{W_1}, \mathbf{W_2}, \mathbf{b_1}, \mathbf{b_2}$ are the parameters for the 2-layer neural network.

After computing the loss, we propagate the gradient back and update the parameters. In our models, we define an unsupervised loss function in Eq.5 which

**Algorithm 1** SiGAT embedding generation (forward) algorithm

```
Input: Sigend directed Graph $G(V, E, s)$;
    Motifs list $\mathcal{M}$; Epochs $T$; Batch Size $B$;
    Motifs graph extract function $F_{m_i}, \forall m_i \in \mathcal{M}$;
    Aggregator GAT-AGGREGATOR$_{m_i}$ with the parameter $\mathbf{W}_{m_i}, \boldsymbol{a}_{m_i}, \forall m_i \in \mathcal{M}$;
    Weight matrices $\mathbf{W_1}, \mathbf{W_2}$ and bias $\mathbf{b_1}, \mathbf{b_2}$;
    Non-linearity function Tanh ;
Output: Node representation $\mathbf{Z}_u, \forall u \in V$
$\mathbf{X}(u) \leftarrow \text{random}(0, 1), \forall u \in V$
$G_{m_i} \leftarrow F_{m_i}(G), \forall m_i \in \mathcal{M}$
$\mathcal{N}_{m_i}(u) \leftarrow \{v|(u, v) \in G_{m_i}\}, \forall m_i \in \mathcal{M}, \forall u \in V$
for $epoch = 1, ..., T$ do
    for $batch = 1, ..., |V|/B$ do
        $\mathcal{B} \leftarrow V_{(batch-1)\times B+1:batch\times B}$
        for $u \in \mathcal{B}$ do
            for $m_i \in \mathcal{M}$ do
                $\mathbf{X}_{m_i}(u) \leftarrow \text{GAT-AGGREGATOR}_{m_i}(\{\mathbf{X}_v, \forall v \in \mathcal{N}_{m_i}(v)\})$
            end for
            $\mathbf{X}'(u) \leftarrow \text{CONCAT}(\mathbf{X}(u), \mathbf{X}_{m_1}(u), ..., \mathbf{X}_{m_{|\mathcal{M}|}}(u))$
            $\mathbf{Z}_u \leftarrow \mathbf{W_2} \cdot \text{Tanh}(\mathbf{W_1} \cdot \mathbf{X}'(u) + \mathbf{b_1}) + \mathbf{b_2}$
        end for
    end for
end for
return $\mathbf{Z}_u$
```

reflects friend embeddings are similar, and the enemy embeddings are dissimilar.

$$J_{\mathcal{G}}(\mathbf{Z}_u) = -\sum_{v^+ \in \mathcal{N}(u)^+} \log\left(\sigma(\mathbf{Z}_u^\top \mathbf{Z}_{v^+})\right) - Q \sum_{v^- \in \mathcal{N}(u)^-} \log\left(\sigma(-\mathbf{Z}_u^\top \mathbf{Z}_{v^-})\right), \quad (5)$$

where $\sigma$ is the sigmoid function, $\mathcal{N}(u)^+$ is the set of positive neighborhoods of node $u$, $\mathcal{N}(u)^-$ is the set of negative neighborhoods of node $u$ and $Q$ is the balanced parameter for the unbalanced positive and negative neighborhoods.

## 4 Experiments

### 4.1 Link Sign Prediction and Dataset Description

Link sign prediction is the task of predicting unobserved signs of existing edges in the test dataset given train dataset. It's the most fundamental signed network analysis task [14].

We do experiments on three real-world signed social network datasets. e.g., Bitcoin-Alpha[3], Slashdot[4] and Epinions[5].

[3] http://snap.stanford.edu/data/soc-sign-bitcoin-alpha.html
[4] http://snap.stanford.edu/data/soc-sign-Slashdot090221.html
[5] http://snap.stanford.edu/data/soc-sign-epinions.html

Bitcoin-Alpha [13] is from a website where users can trade with Bitcoins. Because the Bitcoin accounts are anonymous, the users need to build online trust networks for their safety. Members of Bitcoin-Alpha rate other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1, which can help to prevent transactions with fraudulent and risky users. In this paper, we treat the scores greater than 0 as positive and others as negative. Slashdot [15] is a technology-related news website where users are allowed to tag each other as friends or foes. In other words, it is a common signed social network with friends and enemies labels. Epinions [15] was a product review site. Members of the site can indicate their trust or distrust of the reviews of others. The network reflects people's opinions to others. In these networks, the edges are inherently directed [15].

The statistics of three datasets are summarized in Table 1. We can see that positive and negative links are imbalanced in these datasets.

**Table 1.** Statistics of Three Datasets

| Dataset | # nodes | # pos links | # neg links | % pos ratio |
|---|---|---|---|---|
| Bitcoin-Alpha | 3,783 | 22,650 | 1,536 | 93.65 |
| Slashdot | 82,140 | 425,072 | 124,130 | 77.40 |
| Epinions | 131,828 | 717, 667 | 123,705 | 85.30 |

### 4.2 Baselines

To validate the effectiveness of SiGAT, we compare it with some state-of-the-art baselines and one simple version SiGAT$_{+/-}$. The baselines are as follows:

- Random: It generates $d$ dimensional random values, $Z_j = (z_1, z_2, ..., z_d), z_i \in [0.0, 1.0)$. It can be used to show the logistic regression's ability in this task.
- Unsigned network embedding: We use some classical unsigned network embedding methods (e.g., DeepWalk [17], Node2vec [5], LINE [21]) to validate the importance of signed edges.
- Signed network embedding: We use some signed network embedding methods(e.g., SIDE [10], SiNE [24], SIGNet [9]) to show the effectiveness of modeling signed edges.
- SGCN [2]: It makes a dedicated and principled effort that utilizes balance theory to correctly aggregate and propagate the information across layers of a signed GCN model. It is the latest attempt to introduce GNNs to signed networks.
- FeExtra [14]: This method extracts two parts, a total of 23 features from the signed social network.
- SiGAT$_{+/-}$: We made some simplifications, considering only the positive and negative neighbors, the model can verify if the attention mechanism works.

For a fair comparison, the embedding dimension $d$ of all methods is set to 20 which is closed to FeExtra [14]. We use the authors released code for DeepWalk, Node2vec, LINE, SIDE, SiNE and SIGNet. For SGCN, we use the code from github[6]. We follow the authors recommended settings in Slashdot and Epinions and try to fine-tune parameters to achieve the best results in Bitcoin-Alpha. Like previous works [2,10,24], we first use these methods to get node representations. For edge $e_{ij}$, we concatenate these two learned representation $z_i$ and $z_j$ to compose an edge representation $z_{ij}$. After that, we train a logistic regression classifier on the training set and use it to predict the edge sign in the test set. We perform 5-fold cross-validation experiments to get the average scores as the same as [10]. Each training set is used to train both embedding vectors and logistic regression model. Our models[7] were implemented by PyTorch with the Adam optimizer (Learning Rate = 0.0005, Weight Decay = 0.0001, Batch Size = 500). The epoch was selected from $\{10, 20, 50, 100\}$ by 5-fold cross-validation on the training folds.

### 4.3 Results

**Table 2.** The results of Signed Link Prediction on three datasets

| | | Random | Unsigned Network Embedding | | | Signed Network Embedding | | | Feature Engineering | Graph Neural Network | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Dataset | Metric | Random | Deepwalk | Node2vec | LINE | SiNE | SIDE | SIGNet | FExtra | SGCN | SiGAT$_{+/-}$ | SiGAT |
| Bitcoin-Alpha | Accuracy | 0.9365 | 0.9365 | 0.9274 | 0.9350 | 0.9424 | 0.9369 | 0.9443 | 0.9477 | 0.9351 | 0.9427 | **0.9480** |
| | F1 | 0.9672 | 0.9672 | 0.9623 | 0.9662 | 0.9699 | 0.9673 | 0.9706 | 0.9725 | 0.9658 | 0.9700 | **0.9727** |
| | Macro-F1 | 0.4836 | 0.4836 | 0.5004 | 0.5431 | 0.6683 | 0.5432 | 0.7099 | 0.7069 | 0.6689 | 0.6570 | **0.7138** |
| | AUC | 0.6395 | 0.6435 | 0.7666 | 0.7878 | 0.8788 | 0.7832 | **0.8972** | 0.8887 | 0.8530 | 0.8699 | 0.8942 |
| Slashdot | Accuracy | 0.7740 | 0.7740 | 0.7664 | 0.7638 | 0.8269 | 0.7776 | 0.8391 | 0.8457 | 0.8200 | 0.8331 | **0.8482** |
| | F1 | 0.8726 | 0.8726 | 0.8590 | 0.8655 | 0.8921 | 0.8702 | 0.8984 | **0.9061** | 0.8860 | 0.8959 | 0.9047 |
| | Macro-F1 | 0.4363 | 0.4363 | 0.5887 | 0.4463 | 0.7277 | 0.5469 | 0.7559 | 0.7371 | 0.7294 | 0.7380 | **0.7660** |
| | AUC | 0.5415 | 0.5467 | 0.7622 | 0.5343 | 0.8423 | 0.7627 | 0.8749 | 0.8859 | 0.8440 | 0.8639 | **0.8864** |
| Epinions | Accuracy | 0.8530 | 0.8518 | 0.8600 | 0.8262 | 0.9131 | 0.9186 | 0.9116 | 0.9206 | 0.9092 | 0.9124 | **0.9293** |
| | F1 | 0.9207 | 0.9198 | 0.9212 | 0.9040 | 0.9502 | 0.9533 | 0.9491 | 0.9551 | 0.9472 | 0.9498 | **0.9593** |
| | Macro-F1 | 0.4603 | 0.4714 | 0.6476 | 0.4897 | 0.8054 | 0.8184 | 0.8065 | 0.8075 | 0.8102 | 0.8020 | **0.8449** |
| | AUC | 0.5569 | 0.6232 | 0.8033 | 0.5540 | 0.8882 | 0.8893 | 0.9091 | **0.9421** | 0.8818 | 0.9079 | 0.9333 |

We report the average Accuracy, F1, macro-F1 and AUC in Table 2. We have bolded the highest value of each row. From Table 2, we can find that:

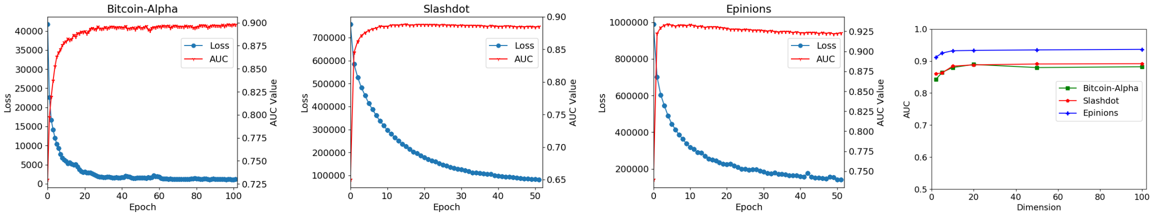
- For signed networks, itself is the positive and negative imbalance. Given random embedding, logistic regression can learn a part of the linear features to a certain extent.
- After using unsigned network embedding methods, all metrics have been increased. Node2Vec has made an obvious improvement, and DeepWalk and LINE have different performances on different datasets.
- SiNE, SIDE, and SIGNet are designed for the signed network; the result is significantly higher than other unsigned network embedding methods. These algorithms modeled some related sociological theories and had achieved good

[6] https://github.com/benedekrozemberczki/SGCN
[7] https://github.com/huangjunjie95/SiGAT

results for signed network analysis. SIGNet showed more stable and better results in our experiments. Besides, The performance of SIDE is not as good as that in [10]. This can be due to the different pre-processing processes and low dimension parameter.

- Although FeExtra was proposed earlier, it has performed well because of its successful use of relevant sociological theory. However, it should be pointed out that this method relies on feature engineering to extract features manually and models only edge features without node representations. Its generalization ability to other tasks is weak.
- SGCN shows a performance close to SiNE as they said in [2], but it cannot effectively model the signed directed networks because the algorithm did not consider the direction. Its mechanism is the simple average of neighbors' hidden state, which can not describe the influence of different nodes to the target node.
- SiGAT$_{\text{only}+/-}$ shows a performance close to or slightly better than SGCN, indicating that the attention mechanism can better model the problem. Compared with the signed graph convolution network method, attention mechanism considers the influence of different nodes on the target node, which is reasonable.
- Our SiGAT model achieves almost the best performance on three datasets. It shows that the relevant theories of sociology have also been successfully modeled by designing different motifs and attention mechanism.

### 4.4 Parameter Analysis



(a) Epoch in Bitcoin-Alpha (b) Epoch in Slashdot (c) Epoch in Epinions (d) Dimension

**Fig. 4.** Parameter Analysis for the epoch and dimension in three datasets

In this subsection, we analyze the some hyper-parameters. We randomly select 80% training edges and 20% testing edges. When analyzing epoch, we set $d = 20$, and record the AUC performance of different epoch generated representations and the total loss value during training. When discussing dimension, we set the corresponding epoch number to 100 (Bitcoin-Alpha), 20 (Slashdot), 10 (Epinions) to discuss the robustness of different dimensions.

As shown in Fig.4, we can see that the value of loss decreases and the value of AUC increases at the earlier epochs and then gradually converges. In different datasets, the best epochs are different. This can be due to the different scales and network structure. Many nodes are updated multiple times quickly due to the mini-batch algorithm. In Fig.4(d), the performance increases quickly and becomes almost stable with the increasing number of dimension $d$. In conclusion, our model shows relatively robust results on hyper-parameters.

## 5 Conclusion

In this paper, we try to introduce GNNs into signed networks. GNNs have recently received widespread attention and achieved great improvements in many tasks. However, most GNNs have been on unsigned networks. We analyzed some key points in signed network analysis and redesigned the GNNs architectures. Based on the basic assumptions of social psychology, we leverage the attention mechanism and different motifs to model the signed network. Experimental results on three real-world signed networks show that our proposed SiGAT outperforms different state-of-the-art signed network embedding methods. Moreover, we analyze the hyper-parameters and show the robustness of our methods. In future work, we will further investigate the performance of SiGAT on more signed network tasks, such as node classification and clustering. Further, we will consider more effective methods to model signed networks.

## Acknowledgement

This work is funded by the National Natural Science Foundation of China under grant numbers 61425016, 61433014, and 91746301. Huawei Shen is also funded by K.C. Wong Education Foundation and the Youth Innovation Promotion Association of the Chinese Academy of Sciences.

## References

1. Arinik, N., Figueiredo, R., Labatut, V.: Signed graph analysis for the interpretation of voting behavior. arXiv preprint arXiv:1712.10157 (2017)
2. Derr, T., Ma, Y., Tang, J.: Signed graph convolutional network. arXiv preprint arXiv:1808.06354 (2018). https://doi.org/10.1109/ICDM.2018.00113
3. Easley, D., Kleinberg, J.: Networks, crowds, and markets: Reasoning about a highly connected world. Cambridge University Press (2010)
4. Gilmer, J., Schoenholz, S.S., Riley, P.F., Vinyals, O., Dahl, G.E.: Neural message passing for quantum chemistry. arXiv preprint arXiv:1704.01212 (2017)
5. Grover, A., Leskovec, J.: node2vec: Scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 855–864. ACM (2016). https://doi.org/10.1145/2939672.2939754
6. Hamilton, W., Ying, Z., Leskovec, J.: Inductive representation learning on large graphs. In: Advances in Neural Information Processing Systems. pp. 1025–1035 (2017)
7. Heider, F.: Attitudes and cognitive organization. The Journal of psychology **21**(1), 107–112 (1946)

8. Hsieh, C.J., Chiang, K.Y., Dhillon, I.S.: Low rank modeling of signed networks. In: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 507–515. ACM (2012). https://doi.org/10.1145/2339530.2339612
9. Islam, M.R., Prakash, B.A., Ramakrishnan, N.: Signet: Scalable embeddings for signed networks. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining. pp. 157–169. Springer (2018). https://doi.org/10.1007/978-3-319-93037-4_13
10. Kim, J., Park, H., Lee, J.E., Kang, U.: Side: representation learning in signed directed networks. In: Proceedings of the 2018 World Wide Web Conference on World Wide Web. pp. 509–518. International World Wide Web Conferences Steering Committee (2018). https://doi.org/10.1145/3178876.3186117
11. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)
12. Kipf, T.N., Welling, M.: Variational graph auto-encoders. arXiv preprint arXiv:1611.07308 (2016)
13. Kumar, S., Spezzano, F., Subrahmanian, V., Faloutsos, C.: Edge weight prediction in weighted signed networks. In: Data Mining (ICDM), 2016 IEEE 16th International Conference on. pp. 221–230. IEEE (2016). https://doi.org/10.1109/ICDM.2016.0033
14. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Predicting positive and negative links in online social networks. In: Proceedings of the 19th international conference on World wide web. pp. 641–650. ACM (2010). https://doi.org/10.1145/1772690.1772756
15. Leskovec, J., Huttenlocher, D., Kleinberg, J.: Signed networks in social media. In: Proceedings of the SIGCHI conference on human factors in computing systems. pp. 1361–1370. ACM (2010). https://doi.org/10.1145/1753326.1753532
16. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U.: Network motifs: simple building blocks of complex networks. Science **298**(5594), 824–827 (2002). https://doi.org/10.1515/9781400841356.217
17. Perozzi, B., Al-Rfou, R., Skiena, S.: Deepwalk: Online learning of social representations. In: Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 701–710. ACM (2014). https://doi.org/10.1145/2623330.2623732
18. Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G.: The graph neural network model. IEEE Transactions on Neural Networks **20**(1), 61–80 (2009). https://doi.org/10.1109/TNN.2008.2005605
19. Schank, T., Wagner, D.: Finding, counting and listing all triangles in large graphs, an experimental study. In: International workshop on experimental and efficient algorithms. pp. 606–609. Springer (2005). https://doi.org/10.1007/11427186_54
20. Sodhani, S., Qu, M., Tang, J.: Attending over triads for learning signed network embedding. Frontiers in Big Data **2**, 6 (2019)
21. Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q.: Line: Large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web. pp. 1067–1077. International World Wide Web Conferences Steering Committee (2015). https://doi.org/10.1145/2736277.2741093
22. Tang, J., Lou, T., Kleinberg, J.: Inferring social ties across heterogenous networks. In: Proceedings of the fifth ACM international conference on Web search and data mining. pp. 743–752. ACM (2012). https://doi.org/10.1145/2124295.2124382
23. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y.: Graph attention networks. arXiv preprint arXiv:1710.10903 **1**(2) (2017)

24. Wang, S., Tang, J., Aggarwal, C., Chang, Y., Liu, H.: Signed network embedding in social media. In: Proceedings of the 2017 SIAM International Conference on Data Mining. pp. 327–335. SIAM (2017). https://doi.org/10.1137/1.9781611974973.37
25. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Yu, P.S.: A comprehensive survey on graph neural networks. arXiv preprint arXiv:1901.00596 (2019)
26. Yuan, S., Wu, X., Xiang, Y.: Sne: signed network embedding. In: Pacific-Asia conference on knowledge discovery and data mining. pp. 183–195. Springer (2017). https://doi.org/10.1007/978-3-319-57529-2_15