2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)
Banff Center, Banff, Canada, October 5-8, 2017

Solving Link-oriented Tasks in Signed Network via an Embedding Approach

Dongfang Du, Hao Wang, Tong Xu*, Yanan Lu, Qi Liu, Enhong Chen
Anhui Province Key Laboratory of Big Data Analysis and Application,
School of Computer Science and Technology, University of Science and Technology of China
{*dfdu, wanghao3, yananlu*}@*mail.ustc.edu.cn*
{*tongxu, qiliuql, cheneh*}@*ustc.edu.cn*

***Abstract*—In this paper, we study the link-oriented tasks in signed network, i.e., labeling link signs and predicting new links. Usually, prior arts directly focus on the link signs, while their intrinsic structural regularities have been largely ignored. Furthermore, these techniques suffer the sensitiveness to the high dimension and sparsity of networks. To deal with these tasks, with verifying the effect of second-order distance in signed network, we propose a novel Link-oriented Signed Network Embedding (LSNE) model, in which network embedding technique is adapted to capture both first-order and second-order distance. Along this line, the link-oriented tasks will be intuitively solved. Extensive experiments on two real-world datasets demonstrate that LSNE could significantly outperform the comparison approaches.**

1. Introduction

Recent years have witnessed the booming of various social network services [18], where the definition of social relationship has been extended as *link signs* emerge. For instance, in Epinions [1], users could select to *trust* or *distrust* other users, which results in **"positive"** or **"negative"** links, and then the so-called *"signed network"* has been formed. This new type of social network raised challenges in link prediction [6] and labeling [15], as in these SNS platforms, many users are not willing to build connections, especially for the negative links, which leads to the severe sparsity of network structure. Thus, identifying hidden relationships, for both links and signs, has been a crucial task.

Traditionally, existing approaches attempt to exploit the structures of signed network in following two aspects. The first ones are usually unsupervised solutions based on topological properties of signed network, like similarity-based [1, 5] and low-rank approximation-based [10, 13, 14] approaches. However, this type of solutions could only fit the undirected signed networks, and the effectiveness of low-rank approximation could be limited even dealing with slightly larger dimension. At the same time, the second ones are always supervised approaches following the idea of classification problems, e.g., they treat degrees, triads and common neighbors as features to train a classifier [7]. However, these approaches are incapable of capturing the intrinsic structural regularities of signed network, and even highly sensitive to the minor alternations of structures. In summary, an effective approach to capture the intrinsic structural regularities is still required for the link-oriented tasks in signed network.

With the development of deep learning techniques, network representations have been widely studied to map nodes in the network into a low-dimension vector space. These techniques offer a more comprehensive perspective to understand the network structure, and further predict future network evolution. With the good performance of DeepWalk [9], which firstly combines random walk with skip-gram model to learning the representations of networks, series of networks embedding methods like LINE [11] and nove2vec [3] have been proposed. However, these works mainly focus on ordinary networks, while the tasks in signed networks still attract less attention.

To deal with these tasks, in this paper, we propose a novel **L**ink-oriented **S**igned **N**etwork **E**mbedding (LSNE) framework. To be specific, two models have benn proposed. The first one, named as *LSNE1*, is inspired by first-order distance and embeds every node into a low-dimension vector. Similarly, the second one, named as *LSNE2*, represents each node as two low-dimension vectors, i.e., output embedding vector and input embedding vector. Different from previous approach, LSNE will not only captures the intrinsic structural regularities, but also consider the direction of signed network. To the best of our knowledge, we are among the first ones to profoundly analyze the relationships between bi-directional embedding vectors. Besides, LSNE enables to exploit embedding vectors as the features for supervised learning and expand to heterogeneous signed networks by conducting this algorithm in every single network and updating alternately.

To the best of our knowledge, we are among the first ones who demonstrate the effect of second-order information in signed network, and then propose a novel framework called LSNE to solve link-oriented tasks in signed network, especially with considering the direction of bi-directional signed network. Experimental results on two real datasets demonstrate that LSNE outperforms several previous methods for these problems and further verifies the proposed approach is capable of capturing the intrinsic structural regularities in signed network.

1. http://www.epinions.com/

2. Signed Network Embedding

In this section, We first provide some definitions of the problems. Then we illustrate the importance of first-order and second-order information in signed network on two public datasets. At last, we introduce our signed network embedding model LSNE and optimization of the model.

2.1. Problem Definition

We first give the definition of the signed network as follows:

DEFINITION 1 (***Signed Network***): The signed network is defined as $G = (U, E^{+}, E^{-})$, where $U = \{u_1, ..., u_n\}$ is the set of vertices, E^{+} and E^{-} respectively represent the sets of positive links and negative links in network.

In signed network, the local pairwise relationship between the nodes is very important. Therefore, the local network structures must be preserved. Different from the definition of the local network structure in general network, we define this in signed network as *First-order Distance* :

DEFINITION 2 (***First-order Distance***): The First-order distance is the local pairwise distance between two nodes. For each node pair (u_i, u_j), the corresponding weight w_{ij} indicates the First-order distance. When the link is positive, like node 1 and node 4 in Figure 1, it indicates that two nodes are similar and their First-order distance is closer. While if the link is negative, it represents they are not similar to each other and their First-order distance is far away, like node 1 and node 8 in Figure 1.

However, only capturing the local structure is not sufficient. Inspired by the balanced theory which means that the friends of friends are friends and the enemies of friends are enemies, we define the *Second-order Distance* as follows:

DEFINITION 3 (***Second-order Distance***): The Second-order distance is determined by the similarity of the neighborhood network structures between a node pair (u_i, u_j). If these two nodes, like node 1 and node 2, share more common nodes with the same sign, then they should be more similar and tend to build a positive link, so their Second-order distance is closer. If these two nodes hold links to more common nodes with different signs, then they prefer to build a negative link and be more dissimilar. Therefore their Second-order distance is far away, for example: node 2 and node 3 in Figure 1.

We investigate both *First-order Distance* and *Second-order Distance* for signed network embedding, which is defined as follows:

DEFINITION 4 (***Signed Network Embedding***): Given a signed network $G = (U, E^{+}, E^{-})$, we **aim** to learn a low-dimensional embedding $\vec{v} \in \mathbb{R}^d$ for each vertex u in signed network, where $d \ll |U|$.

This signed network embedding can preserve *First-order Distance* and *Second-order Distance*. In order to prove this learning embedding's effectiveness, we select the link sign prediction and link prediction tasks to demonstrate it.

Next, we analyze two public datasets to illustrate the importance of *First-order Distance* and *Second-order Distance*, which is the basis of our model.

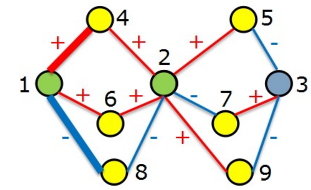

Figure 1. Example of a signed network

TABLE 1. STATISTICS OF THE DATASETS

Datasets	Nodes	Links	Positive Links	Negative Links
Slashdot	77,357	516,575	369,378	120,197
Epinions	131,828	841,372	717,667	123,705

2.2. Data Analysis

In this section, we want to verify the effectiveness of *First-order Distance* and *Second-order Distance*. For the *First-order Distance*, many of the previous research works [2, 5] has proven its importance through matrix decomposition or spectral clustering technologies, etc. Most of them are based on the assumption: *The nodes with positive links are similar and the nodes with negative links are not similar*. We also utilize this as the assumption of our First-order Distance Method. However, for the *Second-order Distance*, a lots of research works have overlooked it or simply utilize it as a part of features. In order to illustrate the importance of the *Second-order Distance*, we utilize two public signed network datasets [7] shown details in Table 1, and put forward two questions as follows:

- If two nodes have more common friends with the same sign links, whether they tend to build a positive link?
- On the contrary, if two nodes have more common friends with the different sign links, whether they tend to build a negative link?

To answer the two questions, we randomly extract 1500 nodes from the two datasets to form a new subgraph for statistical analysis. For arbitrary node pair (u_i, u_j) in subgraph, like u_1 and u_2 in Figure 1, we respectively count the number of their neighbor nodes with the same sign links and different sign links. Then we calculate the ratio of node pair with positive link and negative link, below the different number of neighbor nodes with the same signs and different signs. The final statistical analysis results in the Slashdot Dataset are shown in Figure 2 where the abscissa 10 indicates the number of neighbor nodes is greater than or equal to 10:

From the left figure in Figure 2, we can find that as the number of neighbor nodes with the same sign links increases, the ratio of node pair with positive links significantly increases. This conclusion suggests a positive answer to the first question: *Two nodes with more common friends with the same sign links tend to build a positive link*. For the right figure in Figure 2, we can come to this conclusion that the ratio of node pair with negative links grow with

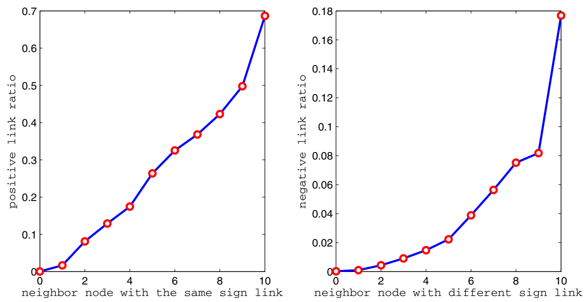

Figure 2. Data analysis in Slashdot

the number of common neighbor nodes with negative links. This answers the second question: *Two nodes with more common friends with the different sign link prefer to build the negative link*. Also the similar results can be observed in Epinions Dataset, we omit it here for brevity.

From the above data analysis, we illustrate the importance of the *Second-order Distance* in signed network. Further more, we make the following assumption that: *The closer the Second-order distance is, the more similar the nodes are, vice versa*, which is the basis of our Second-order Distance Method.

2.3. Link-oriented Signed Network Embedding

In this section, we will propose our methods including first-order distance method and second-order distance method, which capture the intrinsic structures of signed networks.

2.3.1. First-order Distance Method. We extend the first-order distance put forward recently [11] to build our model. The first-order distance preserves the local distance information between two nodes in signed network. Specially, for a positive link (u_i, u_j), we define the joint probability distribution of node u_i and node u_j as :

$$p_1(u_i, u_j) = \frac{1}{1 + exp(-\vec{v}_i^T \cdot \vec{v}_j)}, \tag{1}$$

where $\vec{v}_i \in R^d$ is the low-dimension representation of node u_i, and d is the space dimension given in advance in the experimental setting. Moreover, the empirical distribution of (u_i, u_j) can be defined as $\hat{p}_1(u_i, u_j) = w_{ij}/W^+$ where w_{ij} is the weight of positive link (u_i, u_j), and $W^+ = \sum_{(i,j)\in E^+} w_{ij}$. In order to preserve the first-order distance close for the corresponding nodes of the positive link, an intuitive method is to measure the difference between the two distributions taking advantage of KL divergence:

$$O_1^+ = KL(\hat{p}_1(\cdot,\cdot), p_1(\cdot,\cdot)), \tag{2}$$

By minimizing the KL divergence to keep these distributions close, we can obtain the following formula discarding the irrelevant constants:

$$O_1^+ = -\sum_{(i,j)\in E^+} w_{ij} log p_1(u_i, u_j), \tag{3}$$

For a negative link (u_k, u_s), similarly, we should preserve the fisrt-order distance alienated for the corresponding nodes. Considering that for a positive link (u_i, u_j), it makes $p_1(u_i, u_j)$ increased, which results in the representations of node u_i and node u_j, i.e., $\vec{v}_i$ and $\vec{v}_j$ adjacent. Therefore, if there is also a joint probability distribution $p_2(u_k, u_s)$ for the negative link (u_k, u_s), it will cause $p_2(u_k, u_s)$ increase, so that the representations of these nodes, namely $\vec{v}_k$ and $\vec{v}_s$, will be away from each other. Through the above analysis, we define $p_2(u_k, u_s)$ as follows:

$$p_2(u_k, u_s) = 1 - p_1(u_k, u_s) = \frac{1}{1 + exp(\vec{v}_k^T \cdot \vec{v}_s)}, \tag{4}$$

Equation (4) defines the probability distribution $p_2(.,.)$ in the space of node pairs, while the empirical distribution of (u_k, u_s) can be defined as $\hat{p}_2(u_k, u_s) = w_{ks}/W^-$, where the numerator w_{ks} is the weight of negative link (u_k, u_s), and the denominator $W^- = \sum_{(k,s)\in E^-} w_{ks}$. Similarly, minimizing the KL divergence to obtain the corresponding representations of nodes, after leaving the irrelevant constant, we can get:

$$O_1^- = -\sum_{(k,s)\in E^-} w_{ks} log p_2(u_k, u_s), \tag{5}$$

To sum up, we can get the objective function of first-order distance method in signed networks:

$$O_1 = -(\sum_{(i,j)\in E^+} w_{ij} log p_1(u_i, u_j) + \sum_{(k,s)\in E^-} w_{ks} log p_2(u_k, u_s)), \tag{6}$$

By minimizing the equation (6), we can obtain the representations $\{\vec{v}_i\}_{i=1..|U|}$ in d-dimension space of nodes $\{u_i\}_{i=1..|U|}$.

2.3.2. Second-order Distance Method. We extend the second-order distance put forward recently [11] to build our model, which was verified the effects in signed networks in the data analysis section. The second-order distance method assumes that the nodes with the same sign links to other nodes will become similar, vice versa. For a directed positive link (u_i, u_j) established link from node u_i to node u_j, also inspired by the concept of output matrix and input matrix [17], we define the conditional distribution as:

$$p_3(u_j|u_i) = \frac{exp((\vec{v}_j^{in})^T \cdot \vec{v}_i^{out})}{\sum_{t=1}^{|U|} exp((\vec{v}_t^{in})^T \cdot \vec{v}_i^{out})}, \tag{7}$$

where $|U|$ is the number of node set, $\vec{v}_t^{out}$ is the out-degree representation of node u_t and $\vec{v}_t^{in}$ is the in-degree representation of node u_t. This equation define the conditional distribution $p(.|u_i)$ for every node u_i, which reflects that similar conditional distributions lead to similar nodes. What's more, according to the above mentioned method, this conditional distribution should be as close as possible to the empirical distribution, i.e., $\hat{p}_3(u_j|u_i)$, which can be defined as $\hat{p}_3(u_j|u_i) = w_{ij}/d_i^+$, where w_{ij} is the weight of the positive link (u_i, u_j) and $d_i^+ = \sum_{k\in N^+(i)} w_{ik}$ with $N^+(i)$ means the positive neighbor nodes of node i. Therefore, we leverage KL divergence to minimize the following objective function:

$$O_2^+ = \sum_{u_i \in U} \lambda_i KL(\hat{p}_3(\cdot|u_i), p_3(\cdot|u_i)), \quad (8)$$

where λ_i represents the importance of the node u_i in the signed network, set as d_i^+ for convenience. Further simplify the equation (8), we can obtained:

$$O_2^+ = -\sum_{(i,j)\in E^+} w_{ij} logp_3(u_j|u_i), \quad (9)$$

For a negative link (u_k, u_s) established from node u_k to node u_s, motivated by the relationship between equation (1) and (4), we define the conditional distribution as:

$$p_4(u_s|u_k) = \frac{exp(-(\vec{v}_s^{in})^T \cdot \vec{v}_k^{out})}{\sum_{t=1}^{|U|} exp(-(\vec{v}_t^{in})^T \cdot \vec{v}_k^{out})}, \quad (10)$$

The empirical distribution of $p_4(u_s|u_k)$ can be defined as $\hat{p}_4(u_s|u_k) = w_{ks}/d_k^-$ with the numerator w_{ks} equals to the weight of negative link (u_k, u_s) and the denominator d_k^- equals to $\sum_{t\in N^-(k)} w_{kt}$, where $N^-(k)$ represents the negative neighbor nodes of node k. Similarly, KL divergence is employed to derive this model, then the objective function will be:

$$O_2^- = -\sum_{(k,s)\in E^-} w_{ks} logp_4(u_s|u_k), \quad (11)$$

Through the above analysis, we obtain the objective function in signed networks, utilizing the second-order distance method, namely:

$$O_2 = -(\sum_{(i,j)\in E^+} w_{ij} logp_3(u_j|u_i) + \sum_{(k,s)\in E^-} w_{ks} logp_4(u_s|u_k)), \quad (12)$$

By minimizing the equation (12), we can obtain the representations $\{\vec{v}_i^{out}\}_{i=1..|U|}$ and $\{\vec{v}_i^{in}\}_{i=1..|U|}$ in d-dimension space of nodes $\{u_i\}_{i=1..|U|}$.

2.4. Model Optimization

Since the calculations of the denominator of formula (7) and (10) requires a significant time cost, we utilize negative sampling [8] to optimize the proposed model. Specially, in the second-order distance method, for each positive link (u_i, u_j) established link from node u_i to node u_j, we selected K negative sampling with N positive link (u_i, u_j), and get:

$$N \cdot log\sigma((\vec{v}_j^{in})^T \cdot \vec{v}_i^{out}) + \Sigma_{n=1}^K E_{u_n \sim P_n(u)}[log\sigma(-(\vec{v}_n^{in})^T \cdot \vec{v}_i^{out})], \quad (13)$$

Similarly, for each negative link (u_k, u_s) established link from node u_k to node u_s, we can also get:

$$N \cdot log\sigma(-(\vec{v}_s^{in})^T \cdot \vec{v}_k^{out}) + \Sigma_{n=1}^K E_{u_n \sim P_n(u)}[log\sigma((\vec{v}_n^{in})^T \cdot \vec{v}_k^{out})], \quad (14)$$

where $\sigma(x) = 1/(1+exp(-x))$ is the sigmoid function, the first term in formula (13) represents the observed positive link calculated N times, and the second term represents K negative samples chosen by negative sampling. While in formula (14), the first term represents the observed negative link calculated N times, and the second term represents K negative samples, too. In addition, the probability that each node is sampled is $P_n(u) \propto d_n^{3/4}$, where d_n represents the degree of the node u_n. Besides, considering that the solution of formula (6) will be ordinary, we also employ the idea of negative sampling to solve the first-order distance method. Specifically, for each positive link or negative edge, we just need to replace the out-degree representations and in-degree representations of nodes as the representations of the node in formula (13) or (14) respectively.

3. Experiments

In this section, we conduct extensive experiments to verify the performances of our proposed algorithms with comparsion methods in link-oriented tasks, i.e., sign predict and link predict, followed by the analysis of parameter sensitivities.

3.1. Experimental Settings

3.1.1. Datasets. We utilize the same datasets introduced in the data analysis section to conduct our experiments. Since the number of negative links in the datasets is relatively small, guessing a given link is positive randomly will lead to a high accuracy. Therefore, we refer to the methodology of [4, 7] and obtain two balanced datasets with equal numbers of positive and negative links. Specially, for each negative link, we randomly select a positive link to ensure the eventual datasets for training and testing is balanced.

3.1.2. Comparison Method. Since our model is an unsupervised learning method, some supervised methods like [7] will not be utilized for comparing. Therefore, we select some representative unsupervised algorithms and utilizes Hadamard operator [3] for predictions.

- BAL [16]: This method portrays the balance theory to predict unknown links.
- PMF [10]: This widely used method adopts a probabilistic low-rank matrix factorization and utilizes user-user strength matrix for predictions.
- triMF [13]: This important method extended by PMF ensures to handle directed networks.
- disMF [14]: This method incorporates the balance theory into triMF model. It utilizes balanced information and user-user relationship strength for predictions.
- LINE2 [11]: This method assigns each node a node representation and an auxiliary content representation. It will be extended by our proposed method to suitable for signed networks, only employing node representations for predictions.
- LSNE1: This is the proposed LSNE method with first-order distance method.
- LSNE2: This is the proposed LSNE method with second-order distance method.

3.1.3. Parameters Setup. The parameters of the baselines are set to the optimum values. As for the proposed methods, i.e., LSNE1 and LSNE2, the dimension of representation vector is set as 64. Moreover, the parameters N and K are set from 1, 3, ... , 13 and 1, 3, ... , 9 respectively. The corresponding parameters sensitivity test will be conducted later.

3.1.4. Evaluation Metrics. We leverage two popular metrics, the Accuracy and the Recall, to evaluate the performance of each method in the tasks of sign prediction and link prediction respectively.

3.2. Sign Prediction Task

In this task, we randomly select x% of the links as the set for training models, and further predict the signs of the remaining (100-x)% of links, where x belongs to 20, 40, 60, 80. Our aim is to predict the sign of the testing set after training the model with the training set.

The corresponding experimental results are shown in Figure 3. Our algorithms LSNE1 and LSNE2, outperform the baselines significantly, which verifies that our proposed models are correct. Notice that in slashdot dataset, LSNE1 algorithm performs better than LSNE2, it may because slashdot dataset is more suitable to first-order distance method. And comparing with LINE2, which only utilizes one node representation for predicting, LSNE2 achieves better performance. This phenomenon explains the effectiveness of the method motivated by second-order distance. Moreover, when the training sets become smaller, the models proposed in this paper is more superior to the contrast algorithms, fully showing that the proposed models can significantly deal with sparse data.

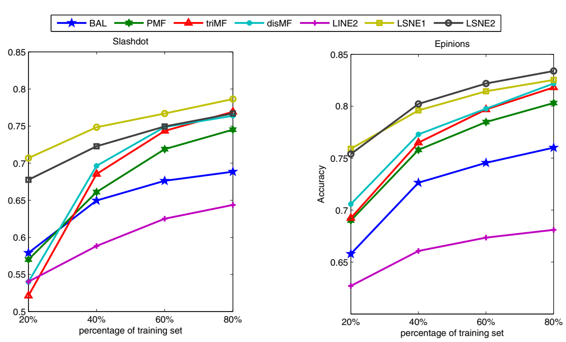

Figure 3. Accuracy in sign prediction task.

3.3. Link Prediction Task

In this task, we intend to predict the TopN positive and negative links, since both of them are significant in recommendations [12] and other applications. Without loss of generality, we choose 80% of the links as a training set, the remaining 20% as a testing set.

Figure 4 and 5 show the Recall@k with different k and our proposed methods significantly outperform the baselines. Comparing with LINE2, our algorithm LSNE2 still achieve better recall, which further demonstrates the importance of out-degree representation and in-degree representation. Besides, we notice that LSNE2 is always better than LSNE1, this may because when data is large enough, second-order distance can capture more information than first-order distance.

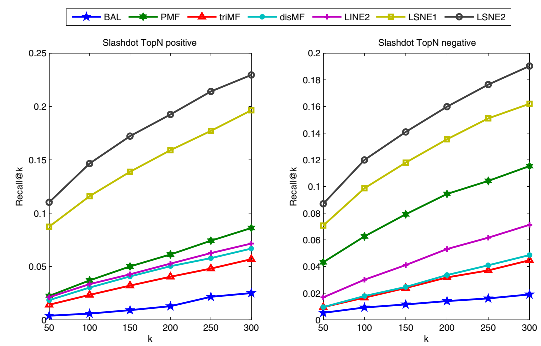

Figure 4. Recall in Slashdot

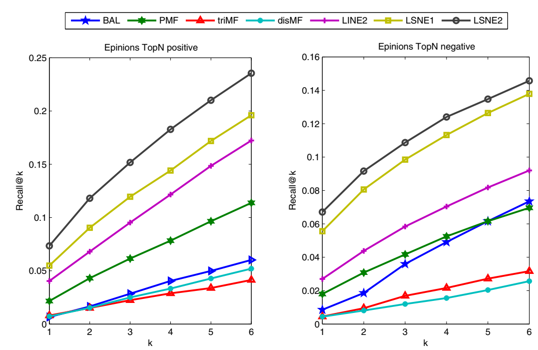

Figure 5. Recall in Epinions

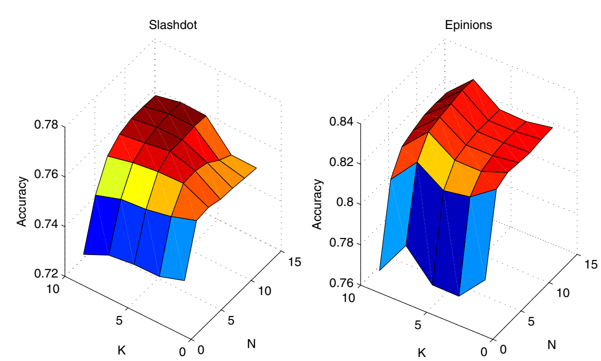

Figure 6. Parameters sensitivity in signed networks

3.4. Parameters Sensitivity

In order the verify the effect of parameter N and K, without loss of generality, we conduct experiments in sign predict task. Figure 6 shows the accuracy of our algorithm LSNE2 in Slashot and Epinions, respectively. We can notice that the accuracy of our algorithm improves with increasing N at the beginning, but improves first and then decreases as K increases. This may because a small amount of negative samples can approximate the original objective function, while a large number of negative samples will cause interference.

4. Conclusion

In this paper, we focus on solving link-oriented tasks in signed networks, and leverage network representation learning method to study them. We propose an effective algorithm, called Link-oriented Signed Network Embedding (LSNE). In this algorithm, we derive first-order distance and second-order distance methods, whose role has been verified in the data statistics phase. The extensive experiments demonstrate the superiority of our proposed algorithm, LSNE. Our future direction is to develop this algorithm and utilize it for heterogeneous signed networks.

5. Acknowledgments

This research was partially supported by grants from the National Natural Science Foundation of China (Grant No. U1605251), the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61325010), and the NSFC Major research program (Grant No. 91546103). Also, this research was supported by the Anhui Provincial Natural Science Foundation (Grant No. 1708085QF140), and the Fundamental Research Funds for the Central Universities (Grant No. WK2150110006).

References

[1] Yi Cen, Rentao Gu, and Yuefeng Ji. Sign inference for dynamic signed networks via dictionary learning. *Journal of Applied Mathematics*, 2013(3):Article ID 708581, 2013.

[2] Rana Forsati, Mehrdad Mahdavi, Mehrnoush Shamsfard, and Mohamed Sarwat. Matrix factorization with explicit trust and distrust side information for improved social recommendation. *Acm Transactions on Information Systems*, 32(4):17, 2014.

[3] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. page 855, 2016.

[4] R. Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *International Conference on World Wide Web*, pages 403–412, 2004.

[5] Cho Jui Hsieh, Kai Yang Chiang, and Inderjit S Dhillon. Low rank modeling of signed networks. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 507–515, 2012.

[6] Ting Jin, Tong Xu, Enhong Chen, Qi Liu, Haiping Ma, Jingsong Lv, and Guoping Hu. Random walk with prefiltering for social link prediction. In *Computational Intelligence and Security (CIS), 2013 9th International Conference on*, pages 139–143. IEEE, 2013.

[7] Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. Predicting positive and negative links in online social networks. In *International Conference on World Wide Web*, pages 641–650, 2010.

[8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26:3111–3119, 2013.

[9] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: online learning of social representations. pages 701–710, 2014.

[10] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *International Conference on Machine Learning*, pages 880–887, 2007.

[11] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. In *International Conference on World Wide Web*, pages 1067–1077, 2015.

[12] Jiliang Tang, Yi Chang, Charu Aggarwal, and Huan Liu. A survey of signed network mining in social media. 49(3), 2015.

[13] Jiliang Tang, Huiji Gao, Xia Hu, and Huan Liu. Exploiting homophily effect for trust prediction. In *ACM International Conference on Web Search and Data Mining*, pages 53–62, 2013.

[14] Jiliang Tang, Xia Hu, and Huan Liu. Is distrust the negation of trust?: the value of distrust in social media. In *ACM Conference on Hypertext and Social Media*, pages 148–157, 2014.

[15] Tong Xu, Dong Liu, Enhong Chen, Huanhuan Cao, and Jilei Tian. Towards annotating media contents through social diffusion analysis. In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, pages 1158–1163. IEEE, 2012.

[16] Shuang Hong Yang, Alexander J. Smola, Bo Long, Hongyuan Zha, and Yi Chang. Friend or frenemy?:predicting signed ties in social networks. In *International Conference on Research on Development in Information Retrieval*, pages 555–564, 2012.

[17] Jihang Ye, Hong Cheng, Zhe Zhu, and Minghua Chen. Predicting positive and negative links in signed social networks by transfer learning. In *International Conference on World Wide Web*, pages 1477–1488, 2013.

[18] Zhou Zhao, Hanqing Lu, Deng Cai, Xiaofei He, and Yueting Zhuang. User preference learning for online social recommendation. *IEEE Transactions on Knowledge and Data Engineering*, 28(9):2522–2534, 2016.