

Link Prediction Algorithm for Signed Social Networks Based on Local and Global Tightness

Miao-Miao Liu*, Qing-Cui Hu*, Jing-Feng Guo**, and Jing Chen**

Abstract

Given that most of the link prediction algorithms for signed social networks can only complete sign prediction, a novel algorithm is proposed aiming to achieve both link prediction and sign prediction in signed networks. Based on the structural balance theory, the local link tightness and global link tightness are defined respectively by using the structural information of paths with the step size of 2 and 3 between the two nodes. Then the total similarity of the node pair can be obtained by combining them. Its absolute value measures the possibility of the two nodes to establish a link, and its sign is the sign prediction result of the predicted link. The effectiveness and correctness of the proposed algorithm are verified on six typical datasets. Comparison and analysis are also carried out with the classical prediction algorithms in signed networks such as CN-Predict, ICN-Predict, and PSNBS (prediction in signed networks based on balance and similarity) using the evaluation indexes like area under the curve (AUC), Precision, improved AUC', improved Accuracy', and so on. Results show that the proposed algorithm achieves good performance in both link prediction and sign prediction, and its accuracy is higher than other algorithms. Moreover, it can achieve a good balance between prediction accuracy and computational complexity.

Keywords

Link Prediction, Sign Prediction, Signed Social Networks, Similarity, Structural Balance Theory, Tightness

1. Introduction

Nowadays, users typically take to online social media to express their opinions, which can inherently be both positive and negative. So, this kind of social networks can be represented by signed social networks. Namely, signed networks are social networks with positive and negative sign attributes. The signs “+” and “-”, respectively indicate the positive relationships (such as support) and the negative relationships (such as opposition) between entities. The research of prediction in signed networks is of great significance for understanding the interaction between positive and negative relations, the formation principle and evolution mechanism of signs, and the analysis of network structure balance. It has important application value in node classification, attitude prediction, recommendation systems in the information field, and the analysis of intermolecular promotion or inhibition in the biological field.

For the prediction research in signed networks, most of the existing algorithms can only complete the

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received November 7, 2020; accepted December 4, 2020.

Corresponding Author: Qing-Cui Hu (hqc70681@163.com)

* Dept. of Computer Science, School of Computer and Information Technology, Northeast Petroleum University, Daqing, Heilongjiang, China (liumiaomiao82@163.com, hqc70681@163.com)

** Dept. of Computer Science and Engineering, College of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei, China (jfguo@ysu.edu.cn, 57178461@qq.com)

prediction of the missing sign. Few of them can achieve the double tasks of link prediction and sign prediction. Though the research of Liu et al. [1] has realized both the sign prediction and link prediction in signed networks, the accuracy of the algorithm depended on the value of the influence factor of the adjustable step size to a certain extent. Besides, relevant researches on mainstream link prediction methods based on the similarity show that, when the step size is greater than 3, the computational complexity of the algorithm increases greatly while its prediction accuracy is not significantly improved. Given the above problems, in order to achieve a better balance between prediction accuracy and computational complexity, a novel algorithm PSSN_TLG (prediction in signed social networks based on tightness of local and global, hereinafter referred to as TLG) is proposed aiming to complete both the link prediction and sign prediction. Based on the structural balance theory, the local link tightness and global link tightness are defined respectively by using the structure information of paths with a step size of 2 and 3 between the two nodes. Then the total similarity is obtained through fusing the above two results to perform the link and sign prediction.

2. Related Work

This paper focuses on the link prediction and sign prediction in undirected signed networks. At present, most of the related works mainly focused on sign prediction. These methods can be roughly divided into supervised learning and unsupervised learning. The prediction method with supervised learning is mainly based on the existing sign attributes of links to select structural features and use decision trees and other classifiers to complete sign prediction. According to the different information used in the selection of network structure features, they are further divided into sign prediction algorithms based on structural balance theory and sign prediction algorithms based on user interaction behavior and context information [2]. For example, in [3], the authors mined the characteristics of unlabeled links and improved the sign prediction accuracy through migration learning of four network features such as node degree, structural balance, status, and potential state. The prediction method with unsupervised learning mainly uses the network structure attributes and the interactive information between nodes to complete sign prediction. They could be roughly divided into two categories, namely, sign prediction based on matrix decomposition or matrix filling and sign prediction based on the node similarity. The representative research achievements are as follows. Su and Song [4] proposed a low-rank matrix decomposition model with offset, which introduced the sign of out-edge and in-edge of neighbor nodes as offset information to improve the accuracy of sign prediction. Shen et al. [5] mainly focused on the prediction of negative links and proposed a framework based on projected non-negative matrix decomposition through unsupervised learning embedded in network structure and user attributes. Although the sign prediction algorithms based on supervised learning and matrix can achieve high prediction accuracy, they usually have high computational complexity, and these models are hard to evaluate. Moreover, due to the poor prediction effect for sparse networks, they are not conducive to the extensive promotion in practical applications. Therefore, the algorithm based on node similarity is still the mainstream link prediction method, and the representative works are as follows. Girdhar et al. [6] proposed a link prediction model based on local and global information, and they distinguished the connection strength between entities based on the fuzzy computing model of trust and distrust. Chen et al. [7] used the set pair theory to treat the sign network as an identity-discrepancy-reverse system and proposed a sign prediction algorithm by

fusing the certainty and uncertainty relations as well as local and global information. Zhu and Ma [8] proposed a highly symmetrical quadrilateral structure by analyzing the sign generation mechanism. Based on the statistical characteristics of local structure, the similarities and divergences of node pairs, and the structural characteristics reflecting the positive and negative attitudes of nodes were extracted. Then the sign prediction was completed.

In addition to the above methods, related scholars have also studied the link prediction in sparse signed networks. For the problem of the sparseness of signed link data in signed networks, i.e., only a small percentage of signed links are given and the number of negative links is much smaller than that of positive links, Beigi et al. [9] proposed a novel signed link prediction model which enabled the empirical exploration of user personalities via social media data. Based on psychology theories, they extracted additional information about users' personalities such as optimism and pessimism that can help determine their propensity in establishing positive and negative links to compensate for data sparsity. Generally, in social networks, users tend to establish positive (or negative) links with those whose generated content we frequently positively (or negatively) interact with online. Based on the verification of these assumptions, a framework for solving the link and interaction polarity prediction problem in signed networks was proposed [10] by understanding the correlation between these two types of opinions from both a local and global perspective. The algorithm was demonstrated to be helpful for both the data sparsity and cold-start problems found in signed networks. In recent years, some scholars have studied the link representation and prediction methods based on convolution neural networks and recurrent neural networks with deep learning mechanism. The ordered node sequence is constructed by using the local topological structure between nodes, and the matrix representation of potential links is generated by node vector expression. Finally, the multi-layer implicit relations of node pairs in the node sequence are extracted based on the neural network operations and used to realize link prediction. However, most of these algorithms focus on link prediction in traditional social networks, and few studies on link prediction and sign prediction for signed networks are found.

3. Problem Statement

3.1 Theoretical Foundation

An undirected and unweighted signed network is usually formalized as $G=(V, E, S)$, in which, $V=\{v_1, v_2, \dots, v_n\}$ represents the node-set, $E=\{e(i, j) \mid v_i, v_j \in V, i \neq j\}$ represents the edge-set, and $S=\{\text{sign}(i, j) \mid v_i, v_j \in V, i \neq j\}$ represents the set of signs with values as follows.

$$e(i, j) = \begin{cases} 1, & (v_i, v_j) \in E \\ 0, & (v_i, v_j) \notin E \end{cases} \quad \text{sign}(i, j) = \begin{cases} +1, & e(i, j)=1 \wedge \text{the link is positive} \\ -1, & e(i, j)=1 \wedge \text{the link is negative} \\ 0, & e(i, j)=1 \wedge \text{the sign is unknown} \vee e(i, j)=0 \end{cases}$$

The structural balance theory provides a theoretical basis for the analysis of undirected signed networks, as shown in Fig. 1. According to this theory, a loop consisting of k edges ($k \geq 3$) is structurally balanced if and only if the product of the sign of all edges is positive. Relevant studies show that the number of balanced rings is far greater than that of unbalanced rings, and with the dynamic development

of the network, unbalanced structures will evolve towards balanced structures. Therefore, the structural balance theory has been widely used in the sign prediction of undirected signed networks.

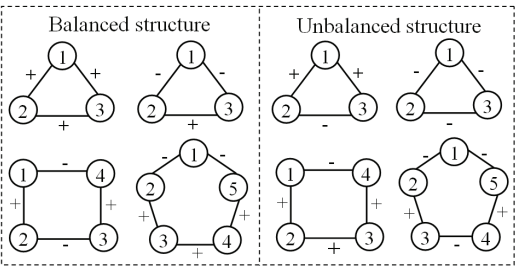


Fig. 1. Sketch of the structural balance theory.

3.2 Problem Definition

The prediction research in signed networks includes two aspects. On the one hand, it is necessary to analyze the probability of the establishment of links between two nodes that have not connected, namely, the link prediction problem. It is generally believed that the higher the similarity between the two nodes, the greater the possibility of establishing a link between them. On the other hand, it is necessary to analyze the sign type of new links or existing links without sign attributes, namely, the sign prediction problem. Generally, the algorithm should strive to ensure that the ring where the predicted link is located can maximize the structural balance of the network as much as possible. Therefore, in the paper, we define the similarity between two nodes based on the characteristics of network topology and the structural balance theory aiming to achieve the two tasks of link prediction and sign prediction.

The goal of the algorithm can be expressed as follows. Given $G=(V, E, S), \forall v_i, v_j \in V$, we aim to predict the possibility of establishing a link between the node v_i and v_j where $e(i, j)=0$ and the corresponding sign type of the link. The algorithm also predicts the missing sign attribute of existing links where $e(i, j)=1$. As shown in Fig. 2, the solid line represents the known edge, and the dotted line represents the unknown edge (which does not exist or exists but has not yet been observed). Nodes of v_1 and v_{10} are not directly connected at present, but due to the existence of common neighbors, v_1 can reach v_{10} through multiple paths. That is to say, the probability of establishing a link between them is very high. However, due to the prediction results of $\text{sign}(1,10)$ based on each balanced ring are different which can be positive or negative, the algorithm should effectively integrate the local and global structural features that affect the similarity of the two nodes to define the total similarity based on the concept of the k -balanced ring. And its ultimate goal is to get the probability of establishing a link between v_1 and v_{10} and to predict the corresponding sign type, namely, $\text{sign}(1,10)$.

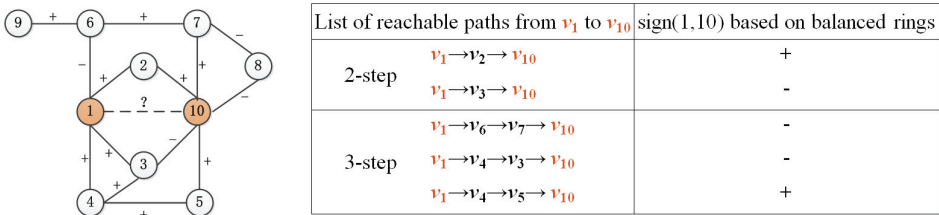


Fig. 2. Diagram of problem definition of the TLG algorithm.

4. Proposed Algorithm

4.1 Core Ideas and Related Definitions

According to the local structure information of the social network graph, the classical CN-index took the number of first-order common neighbors of the node pair into account, and the Jaccard index considered the number of non-common neighbor nodes. While the AA index and RA index considered the degree of the first-order common neighbors. The four indicators respectively focused on one of the factors that affect the similarity, which ignored the influence of other local features on the similarity of the node pair. Herein, using $\text{sim}(x,y)$ represents the similarity value of the node pair $\langle v_x, v_y \rangle$. According to CN and Jaccard indexes, $\text{sim}(1,10)$ in Fig. 3(a) is greater than $\text{sim}(1,10)$ in Fig. 3(b). According to AA and RA indexes, $\text{sim}(1,10)$ in Fig. 3(a) is greater than $\text{sim}(1,10)$ in Fig. 3(c).

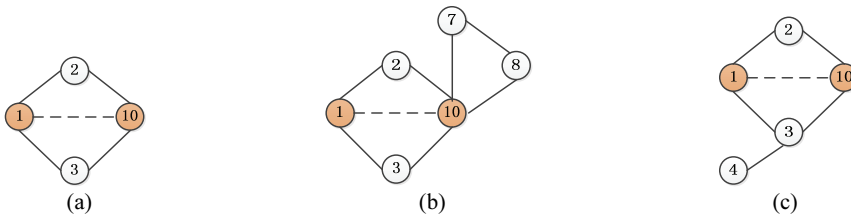


Fig. 3. Diagram of similarity definition based on local tightness.

Definition 4.1 (Local link tightness). To improve the prediction accuracy, the TLG algorithm considers the degree of the two nodes, their first-order common neighbors and non-common neighbors, the sign of the edge, and other local structural characteristics. $\forall v_x, v_y \in V$ and $\text{sign}(x,y)=0$, based on the concept of the structural balanced ring, the local link tightness of the node pair is defined by using the information of paths connecting the two nodes with a step size of 2, which is recorded as $\text{TLSim}(x, y)$. In this paper, $k(x)$ represents the degree of the node v_x , $N_1(x)$ and $N_2(x)$, respectively, represent the first-order and second-order neighbor sets of node v_x , and $|N_1(x)|$ represents the size of the set $N_1(x)$.

$$\text{TLSim}(x, y) = \frac{\sum_{z \in N_1(x) \cap N_1(y)} \frac{\text{sign}(x, z) \times \text{sign}(z, y)}{k(z)}}{|N_1(x) \cup N_1(y)|} \quad (1)$$

Definition 4.2 (Global link tightness). According to the global structure information of the network, the LP and Katz indexes consider the influence of the path with a step size of 2 and 3 or all paths on the similarity respectively and assign weight decay factors to paths with different step size. But the LP index ignores the impact of local information on the similarity. While Katz index needs to calculate all path information, which will lead to high complexity. Given this, to achieve a better balance between accuracy and complexity, the TLG algorithm takes the influence of structure information of path with a step size of 3 connecting the two nodes on the similarity into account. Firstly, the link strength of two directly connected nodes is defined. Then, based on the concept of the structural balanced ring, the global link tightness is defined by the link strength of three groups of intermediate transmission node pairs on the path that connects the two nodes with the step size of 3, written as $\text{TGSim}(x, y)$.

$$\text{TGSim}(x,y)=\sum_{l_k}\frac{\text{sign}(x,z_k')}{\log k(x)+\log k(z_k')+1}\times\frac{\text{sign}(z_k',z_k'')}{\log k(z_k')+\log k(z_k'')+1}\times\frac{\text{sign}(z_k'',y)}{\log k(z_k'')+\log k(y)+1}\quad (2)$$

In the definition of $\text{TGSim}(x,y)$, l_k is the K th path with a step size of 3 connecting v_x and v_y , namely, $l_k = v_x e(x, z_k') v_{z_k'} e(z_k', z_k'') v_{z_k''} e(z_k'', y) v_y$. While $v_{z_k'}$ and $v_{z_k''}$ are the two intermediate nodes on the path l_k . That is to say, $v_{z_k'} \in N_1(x) \cap N_2(y)$, $v_{z_k''} \in N_1(y) \cap N_2(x)$, $e(x, z_k')=1$, $e(z_k', z_k'')=1$, $e(z_k'', y)=1$.

Definition 4.3 (Total similarity between the two nodes). The total similarity between two unconnected nodes is defined as the sum of the local and global link tightness of the two nodes, which is denoted as $\text{TLGSim}(x, y)$, as follows. $|\text{TLGSim}(x, y)|$ represents the possibility of the node v_x and v_y to establish a link, and the prediction result of $\text{sign}(x, y)$ is the same as the sign type of $\text{TLGSim}(x, y)$.

$$\text{TLGSim}(x, y) = \text{TLSim}(x, y) + \text{TGSim}(x, y) \quad (3)$$

4.2 Description and Implementation

Algorithm TLG

Input: $G = (V, E, S)$, which describes the static snapshot graph of an undirected signed network.

Output: the Top- k links that are the most likely to establish in G , their corresponding signs, and the missing sign attributes of existing edges in G .

$\forall v_x, v_y \in V$, if $e(x, y)=0$ or $e(x, y)=1$ but $\text{sign}(x, y)=0$, perform the following operations.

- 1) Read the dataset file and obtain the adjacency matrix and adjacency table of the graph G .
- 2) Find all of the node pair $\langle v_x, v_y \rangle$, get the adjacency lists of v_x and v_y , and calculate $\text{TLSim}(x, y)$.
- 3) Get all the second-order neighbors of v_x and v_y , and calculate $\text{TGSim}(x, y)$.
- 4) Calculate $\text{TLGSim}(x, y)$.
- 5) Do the prediction. In terms of the prediction of the missing sign of the existing edge, if $\text{TLGSim}(x, y) > 0$, then $\text{sign}(x, y) = +1$. While if $\text{TLGSim}(x, y) < 0$, $\text{sign}(x, y) = -1$. In terms of the prediction of unknown links, we sort $|\text{TLGSim}(x, y)|$ in descending order, and output the node pairs corresponding to the Top- k values, which have the highest probability of establishing links. And the prediction result of $\text{sign}(x, y)$ corresponding to the unknown link is the same as the sign type of the value of $\text{TLGSim}(x, y)$.

5. Experiments

Six representative datasets are selected, and the training set and test set are divided by the 10-fold cross-validation method. The improved AUC' and Accuracy' are used as evaluation indexes. And the results of the proposed algorithm are compared with CN (common neighbor), ICN (improved common neighbor), and PSNBS (prediction in signed networks based on balance and similarity) algorithm [1].

5.1 Dataset

Three classical large-scale real datasets and three small datasets are used for the experiment. The basic information is shown in Table 1. Among them, the FEC and CRA network are two simulation datasets,

and the Gahuku Gama Sub-tribes signed network, written as GGS, is a real dataset.

Table 1. Basic characteristics of the dataset

Dataset	V	E	Percentage of edges (%)	
			Positive	Negative
Epinions	131828	840799	85.00	15.00
Slashdot	82144	549202	77.40	22.60
Wikipedia	138592	740106	78.70	21.30
FEC	28	42	71.40	28.60
GGS	16	58	50.00	50.00
CRA	36	74	93.20	6.800

5.2 Evaluation Index

The classical index AUC [2] is suitable for the evaluation of link prediction accuracy in traditional social networks with only positive links. However, the total similarity score between two nodes in the TLG algorithm may be positive or negative. Its absolute value represents the probability of the two nodes to establish a link, and it is positive or negative represents the sign type of the predicted link. Therefore, in our experiments, the classical AUC index is adjusted to obtain a new index, written as AUC', which can correctly evaluate the link prediction results in signed networks. Its definition is shown as follows. In the experiment, the total similarity of each node pair corresponding to the edge selected randomly from the test set and the edge selected from the non-existent edge-set is calculated. And the comparison is done only when the signs of the two similarity values are the same. If the absolute value of the total similarity between the two nodes corresponding to the edge selected from the test set is greater than that of the edge selected from the nonexistent edge-set, $\bar{n}'$ will plus 1. If the absolute values of the two are equal, $\bar{n}''$ will plus 1. If the signs of the two are different, the calculation will be canceled, and we will choose edges again. In each group of experiments, it should be ensured that there are 20,000 times that the signs of the total similarity value corresponding to the two edges selected are the same. That is to say, the value of $\bar{n}$ is 20,000.

$$AUC' = \frac{\bar{n}' + 0.5 \times \bar{n}''}{\bar{n}} \quad (4)$$

For the sign prediction in signed networks, a series of evaluation indexes can be obtained, such as TP, FP, TN, FN, Recall, Precision, Accuracy, F1-score, etc. The sign prediction in signed networks needs to evaluate the comprehensive prediction accuracy of positive and negative signs. Relevant researches have shown that in most real signed networks the ratio of the number of positive links to the number of negative links is greater than 3:1. That is to say, the probability that positive links are selected in the experiment is much higher than that of negative links. Therefore, we improved the index of Accuracy, and the index Accuracy' is obtained by adjusting the above indexes to give a weight of 1 for sign prediction results of positive edges and a weight of 0.5 for sign prediction results of negative edges, which is used as the comprehensive evaluation index, as follows.

$$Accuracy' = \frac{TP + 0.5 \times TN}{(TP + FN) + 0.5 \times (TN + FP)} \quad (5)$$

5.3 Experimental Results and Analysis

(1) Analysis of link prediction accuracy: The experimental results of the link prediction accuracy of the TLG algorithm based on the AUC' evaluation index are shown in Fig. 4. Prediction accuracy and the average value of ten independent experiments which is abbreviated as AVG are obtained. The number of links that are effectively predicted and compared in each independent experiment is 20,000, namely, $\overline{n} = 20,000$. As can be seen from Fig. 4, the TLG algorithm shows high link prediction accuracy for the three large real signed networks and the small simulation network CRA (as shown in Figs. 5 and 6). These four datasets are signed networks with conventional topology in which the number of positive links is much larger than that of negative links and the degree distribution of nodes has no particularity.

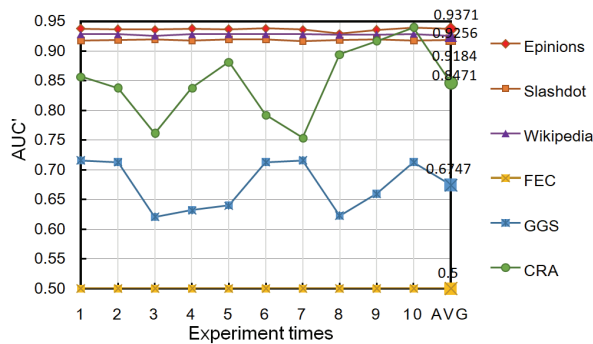


Fig. 4. Prediction results of TLG based on AUC'.

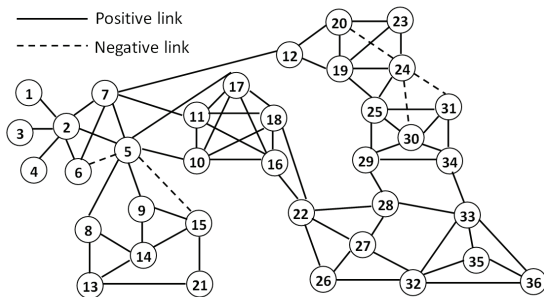


Fig. 5. Topology of CRA network.

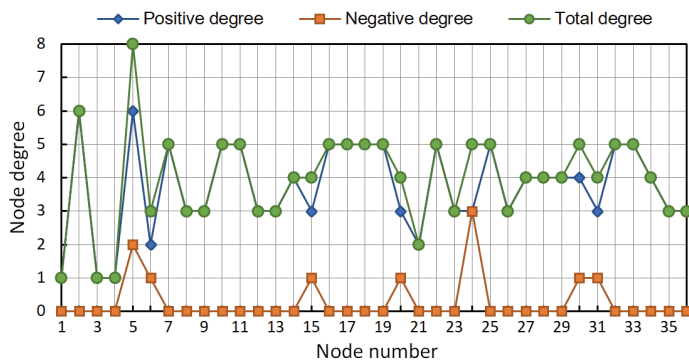


Fig. 6. Degree distribution of CRA network.

For the GGS network, the accuracy of the proposed algorithm is lower than that of the first four datasets. This network describes the real political alliance and enmity among 16 sub-tribes as shown in Fig. 7. The proportion of positive links and negative links is 1:1, and the degree distribution of 16 nodes is shown in Fig. 8. For the dataset with the same number of positive and negative links, the prediction accuracy of the TLG algorithm can still reach 67%, which shows its good robustness.

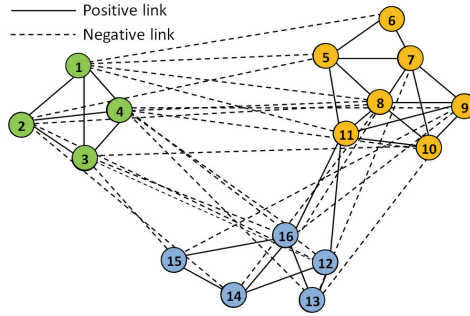


Fig. 7. Topology of GGS network.

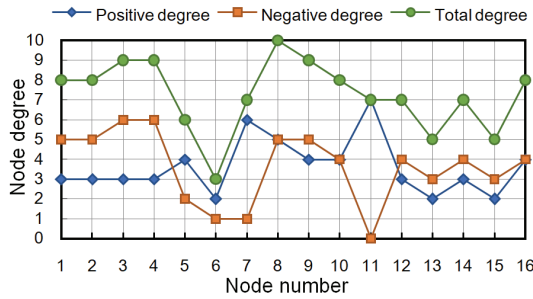


Fig. 8. Degree distribution of GGS network.

For the FEC network, the accuracy of the TLG algorithm remains 0.5. The topology and node degree distribution of the dataset are shown in Figs. 9 and 10. It can be seen that 28 nodes are divided into two types with the same degree distribution. When calculating AUC' , the topological structure of the links obtained from the test set and the nonexistent edge set is almost the same in most cases. In other words, the probability that the degree distribution of the node pairs corresponding to these two links is not the same is $C_{24}^2 C_4^2 / C_{28}^2 C_{26}^2$, which equals 0.0135. Namely, $\bar{n}' \approx 0$ and $\bar{n}'' \approx \bar{n}$. Therefore, the value of AUC' should be 0.5. The experimental results also verify the correctness of the TLG algorithm.

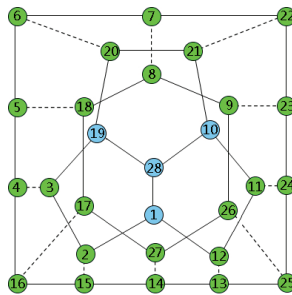


Fig. 9. Topology of FEC network.

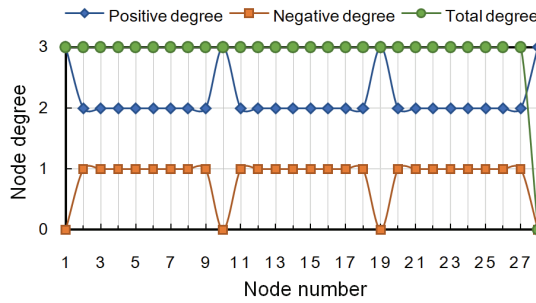


Fig. 10. Degree distribution of FEC network.

(2) Analysis of sign prediction accuracy: Similarly, the sign prediction accuracy of the proposed algorithm is verified on six datasets. The experimental results are shown in Table 2 and Fig. 11.

Table 2. Experimental results of sign prediction of TLG

	Epinions	Slashdot	Wikipedia	FEC	GGS	CRA
TP(++)	0.8218	0.6867	0.8659	0.6000	0.5000	0.8750
FP(+/-)	0.0623	0.0989	0.0617	0.2000	0.1667	0.0000
TN(-v)	0.1048	0.1511	0.0593	0.0000	0.3333	0.0000
FN(-/+)	0.0111	0.0633	0.0131	0.2000	0.0000	0.1250
Recall	0.9866	0.9156	0.9851	0.7500	1.0000	0.8750
Precision	0.9295	0.8741	0.9335	0.7500	0.7500	1.0000
F1-score	0.9573	0.8944	0.9586	0.7500	0.8571	0.9333
Accuracy	0.9266	0.8378	0.9252	0.6000	0.8333	0.8750

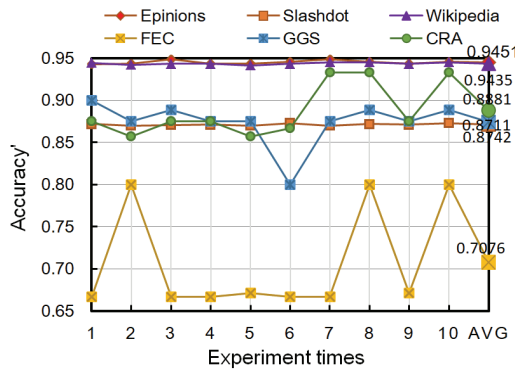


Fig. 11. Prediction results of TLG based on Accuracy'.

(3) Comparison with other algorithms: Experimental comparisons are carried out with the CN-Predict, ICN-Predict and the PSNBS algorithm. The AUC', Accuracy', and AUC [1] are used as evaluation indexes. Results are shown in Fig. 12, Fig. 13, and Fig. 14, respectively. Note that the results of the PSNBS algorithm are the highest accuracy when the step size influence factor λ is assigned the optimal value. Results show that the TLG algorithm does not need to preset the influence factor, and it can achieve a good prediction effect for both conventional large-scale signed networks and small-scale signed networks with the special topology. Moreover, its accuracy is higher than other algorithms.

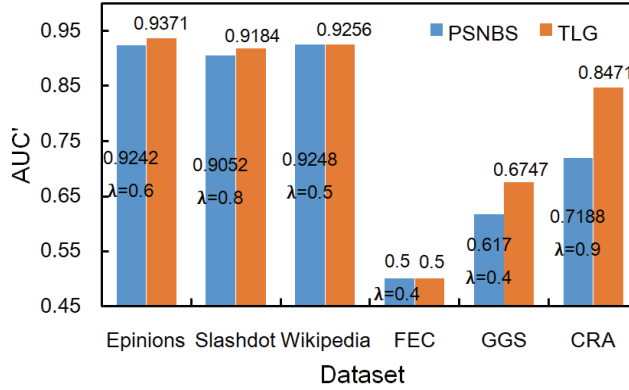


Fig. 12. Link prediction results based on AUC'.

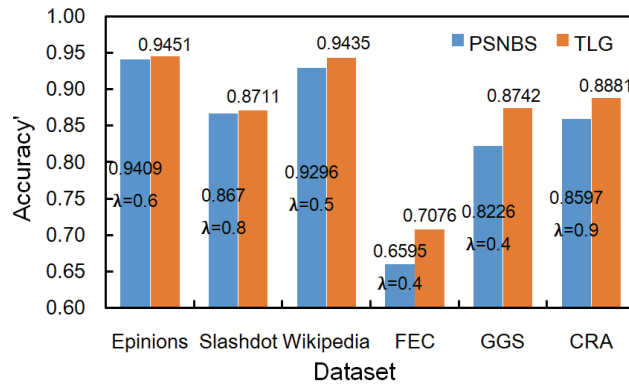


Fig. 13. Sign prediction results based on Accuracy'.

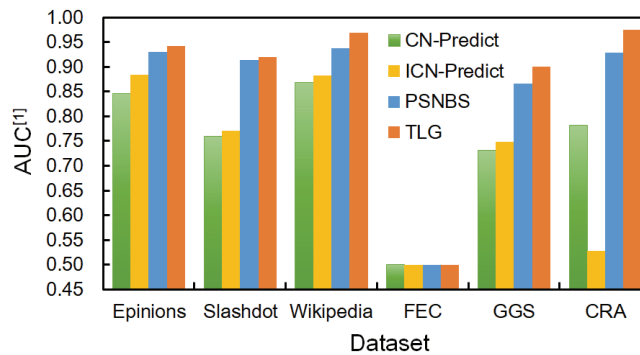


Fig. 14. Comparison results based on AUC [1].

(4) The Top- k recommended links: According to the TLG algorithm, the Top-100 recommended links are given for the first three large-scale networks, as shown in Figs. 15–17. Meanwhile, the top-50 recommended links are also given for the two small-scale datasets, as shown in Fig. 18. In these four figures, the serial number of the node pairs and their similarities are displayed. And the prediction of positive links and negative links are respectively represented by different colors.

Besides, the recommendation results of links based on the TLG and PSNBS algorithm are compared

on the two small datasets, as shown in Fig. 19. Results of the proposed algorithm are consistent with that of the PSNBS algorithm, which further verifies the correctness and effectiveness of the TLG algorithm.

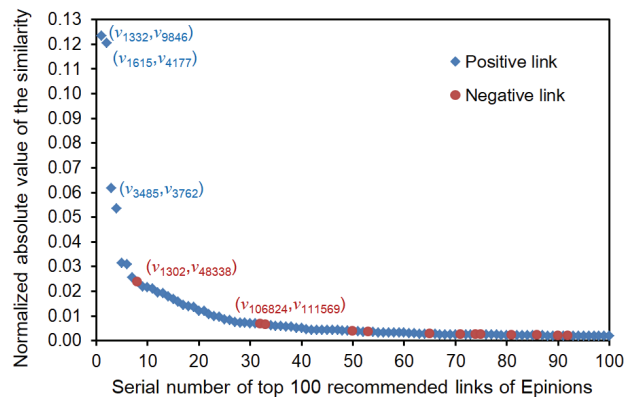


Fig. 15. Recommended links for Epinions.

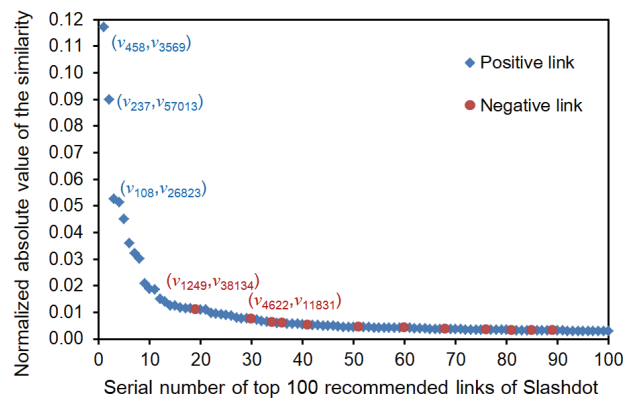


Fig. 16. Recommended links for Slashdot.

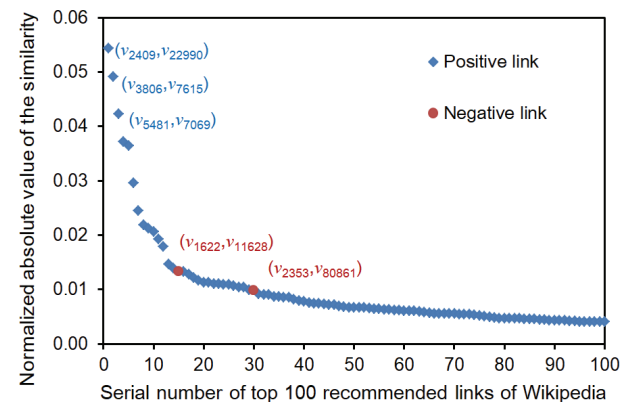


Fig. 17. Recommended links for Wikipedia.

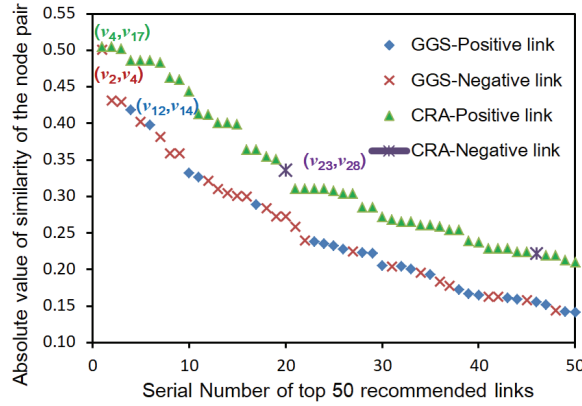


Fig. 18. Recommended links for GGS and CRA

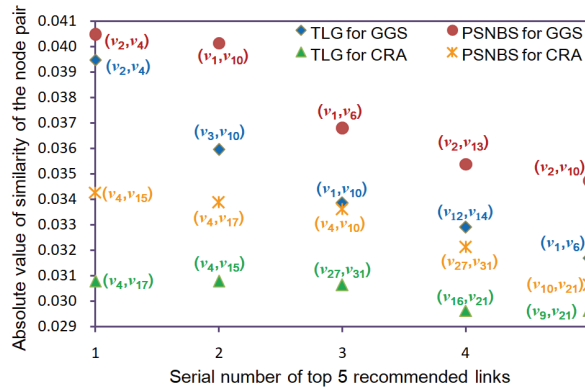


Fig. 19. Comparison of results of recommended links.

6. Conclusion

A link prediction algorithm for signed networks is proposed. It effectively obtains the local and global structural features that affect the similarity of nodes and can achieve the dual goals of link prediction and sign prediction. Experimental results show its better prediction effect than other methods. However, for large-scale signed networks, how to reduce the computational complexity is a key problem. Besides, the proposed algorithm takes the static snapshot of the signed network as the research object while the real social network is dynamic. Therefore, in the next research work, perhaps we should integrate deep learning methods to mine more structural properties to complete the link prediction of signed networks.

Acknowledgement

This work was supported by the Natural Science Foundation of China (No. 42002138 and 61871465), Natural Science Foundation of Heilongjiang Province (No. LH2019F042 and LH2020F003), Youth Science Foundation of Northeast Petroleum University (No. 2018QNQ-01), and Science & Technology Program of Hebei (No. 20310301D).

References

- [1] M. M. Liu, J. F. Guo, and J. Chen, "Link prediction in signed networks based on the similarity and structural balance theory," *Journal of Information Hiding and multimedia Signal Processing*, vol. 8, no. 4, pp. 831-846, 2017.
- [2] M. M. Liu, Q. C. Hu, J. F. Guo, and J. Chen, "Survey of link prediction algorithms in signed networks," *Computer Science*, vol. 47, no. 2, pp. 21-30, 2020.
- [3] D. Li, D. Shen, Y. Kou, Y. Shao, T. Nie, and R. Mao, "Exploiting unlabeled ties for link prediction in incomplete signed networks," in *Proceedings of 2019 3rd IEEE International Conference on Robotic Computing (IRC)*, Naples, Italy, 2019, pp. 538-543.
- [4] X. Su and Y. Song, "Local labeling features and a prediction method for a signed network," *CAAI Transactions on Intelligent Systems*, vol. 13, no. 3, pp. 437-444, 2018.
- [5] P. Shen, S. Liu, Y. Wang, and L. Han, "Unsupervised negative link prediction in signed social networks," *Mathematical Problems in Engineering*, vol. 2019, article no. 7348301, 2019. <https://doi.org/10.1155/2019/7348301>
- [6] N. Girdhar, S. Minz, and K. K. Bharadwaj, "Link prediction in signed social networks based on fuzzy computational model of trust and distrust," *Soft Computing*, vol. 23, no. 22, pp. 12123-12138, 2019.
- [7] X. Chen, J. F. Guo, X. Pan, and C. Zhang, "Link prediction in signed networks based on connection degree," *Journal of Ambient Intelligence and Humanized Computing*, vol. 10, no. 5, pp. 1747-1757, 2019.
- [8] X. Zhu and Y. Ma, "Sign prediction on social networks based nodal features," *Complexity*, vol. 2020, article no. 4353567, 2020. <https://doi.org/10.1155/2020/4353567>
- [9] G. Beigi, S. Ranganath, and H. Liu, "Signed link prediction with sparse data: the role of personality information," in *Companion Proceedings of the 2019 World Wide Web Conference*, San Francisco, CA, 2019, pp. 1270-1278. <https://doi.org/10.1145/3308560.3316469>
- [10] T. Derr, Z. Wang, J. Dacon, and J. Tang, "Link and interaction polarity predictions in signed networks," *Social Network Analysis and Mining*, vol. 10, article no. 18, 2020. <https://doi.org/10.1007/s13278-020-0630-6>



Miao-Miao Liu <https://orcid.org/0000-0002-1667-9519>

She was born in 1982 and she is currently a full professor and the Master's Supervisor of Northeast Petroleum University in China. She got her doctorate in computer science and technology at Yanshan University in 2017. Her main research interests include community discovery and link prediction in social networks.



Qing-Cui Hu <https://orcid.org/0000-0001-8440-0703>

She was born in 1982 and she is currently a full professor and the Master's Supervisor of Northeast Petroleum University in China. She got her doctorate in computer science and technology at Yanshan University in 2017. Her main research interests include community discovery and link prediction in social networks.



Jing-Feng Guo <https://orcid.org/0000-0003-2373-6523>

He is currently a full professor and Doctoral supervisor of Yanshan University in China. His research interests include database theory, data mining and social network analysis.



Jing Chen <https://orcid.org/0000-0003-2085-5877>

She is currently an associate professor and Master supervisor of Yanshan University in China. Her research interests include community discovery and information dissemination in social networks.