Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **4**, July **2016**

Link Prediction Based on The Similarity of Transmission Nodes of Multiple Paths in Weighted Social Networks

Miaomiao Liu

Northeast Petroleum University
Daqing, Heilongjiang 163318, China
liumiaomiao82@163.com

Jingfeng Guo and Xu Luo

College of Information Science and Engineering
Yanshan University
Qinhuangdao, Hebei 066004, China

Received November, 2015; revised April, 2016

Abstract. *Most link prediction algorithms only consider local or global characteristics of the graph, so it is difficult to reach equilibrium in the accuracy and the computational complexity. And the research on link prediction in weighted networks is relatively less. A new algorithm STNMP (Similarity based on Transmission Nodes of Multiple Paths) for link prediction in weighted social networks is proposed. Firstly, the concept of the edge weight strength is introduced to measure the local similarity of neighbor node pairs. Then the definition of the path similarity contribution and the similarity of transmission nodes of multiple paths are given which are used to describe the total contribution of all these paths of 2 and 3 paces to the similarity of node pairs. The effectiveness of the algorithm is verified through experiments on many real networks using AUC and Precision as evaluation index. Furthermore, the prediction accuracy and efficiency of the algorithm are also analyzed through the comparison with those classical link prediction algorithms based on the similarity index of CN, AA, etc. The results show that for the small scale of social networks, the accuracy of STNMP algorithm is higher than those of existing algorithms. In addition, with a good generality, the algorithm is also applicable to the link prediction in unweighted social networks.*
Keywords: Link prediction; Weighted social networks; Path similarity contribution; Transmission nodes; Multiple paths.

1. **Introduction.** As social networks are highly dynamic and relationships between the entities in the network are always developing and evolving, link prediction has become a hot study with a wide application in the information retrieval, recommendation system, analysis on dynamic evolution of structure of social networks [1], classification of nodes in signed networks [2], etc. Link prediction refers to the estimation of the possible existence of a link between two nodes based on the observed links and the attributes of the nodes, which can be categorized into the unknown link prediction and the future link prediction [3]. The former is the prediction of the links that have not yet been found, though they have come into existence. As a process of data mining, it has great research significance and wide application in the protein interaction networks [4] and many other biological networks. The latter is the prediction of the links that may appear in the future which

is relevant with the evolution of the network. This paper focuses on the research of the future link prediction.

2. **Related Work.** There has been a lot of research on link prediction of social networks such as algorithms based on Markov chains [5-6] or machine learning [7], maximum likelihood estimation [8-9], the probability models [10] and the similarity. [3] gave a review and comparison of several representative link prediction methods. The mainstream link prediction methods are algorithms using the essential attributes of nodes to define the node similarity. That is to say, two nodes are considered to be more similar if they have many common features. CN algorithm [11] considered that if two nodes have more common neighbors, it would be more likely to form a link between them. Jaccard algorithm [12] defined the similarity of node pairs as the ratio of the union and the intersection of their neighbor nodes. Adamic-Adar algorithm [13] considered the degree of the neighbors of node pairs. It held that nodes with smaller degree had greater influence on the similarity than nodes with larger degree. PA algorithm [14] defined the similarity of node pairs as the product of the degree of these two nodes. This method was less complicated in calculation, but the prediction accuracy was not high. The shortest path algorithm [15] defined the similarity as the minimum of the path length between two nodes. Katz algorithm [16] considered the influence of all paths and paths of different step-size on the similarity of node pairs, thus it is more complicated in calculation. LP [17] algorithm only considered the direct neighbors and indirect neighbors so as to reach a compromise on the prediction accuracy and computational complexity. According to the SimRank algorithm [18], the existence of the link between two nodes was relevant with adjacent links of it. [19] studied 9 famous local similarity index, meanwhile it put forward two kinds of new local index. The FriendTNS algorithm [20] used the product of the local similarity of transitive nodes on the shortest path to measure the extended similarity. The CNBIEC algorithm [21] improved the prediction accuracy by using the information of links between common neighbor nodes. [22] put forward a link prediction algorithm based on the dependent degree of two links, and it focused on the research of the relationship of these links. The link prediction for weighted networks is very important, however the related systematic research is still less. [23] used the method based on weighted similarity measurement to predict links and obtained good results in the dense social networks. [24] used the local similarity index to estimate the possible existence of links in weighted networks. It also put forward three kinds of weighted similarity indices which can be considered as avariants of CN, AA and RA respectively. But the experimental results of these weighted index in the Net Science network and US Airports network were not ideal. [25] discussed how to use the similarity to predict the type of the node in some labeled networks. [26] proposed a random walk-based time link prediction algorithm in weighted networks by regarding the commonly concerned topics between users in recent time as the weights of the edge.

Most existing link prediction algorithms based on the similarity are aiming at the unweighted networks, and they can be divided into two categories. One is the similarity algorithms based on local information of nodes such as CN, AA which mainly use the information of neighbor nodes degree. Simple and easily implemented, these algorithms obtain better prediction results. But they ignore the links between neighbor nodes so they can not effectively explore the influence of network topology on the node similarity. The other is similarity algorithms based on the path structure such as Katz index which consider all paths contribution to the node similarity. However they ignore the local similarity of nodes on the path and the calculation of path information is complex.

In view of the limitations of the existing algorithms, a new link prediction algorithm STNMP is proposed. It can achieve link prediction in weighted networks based on multiple paths and the local similarity of adjacent transmission nodes on these paths. Section 3 is the main idea of the algorithm. Section 4 is preliminaries which give the related concepts and the definition of similarity index. Section 5 is the implementation steps of the algorithm. Section 6 is experiments and analysis. The last section is the conclusion.

3. **Main Ideas of the STNMP Algorithm.** According to the link prediction algorithms based on the similarity, if two nodes have higher similarity, they would have greater possibility of establishing links. The key of the algorithm is effectively capturing the effect of local and global features of the network on the node similarity and giving the reasonable formula definition for the calculation of the similarity so as to improve the prediction accuracy and efficiency.

When measuring the effect of the node′s local properties on the node similarity, we think that if the two nodes are not neighbor, their local similarity is 0. Nodes with smaller degree contribute more to the local similarity than nodes with larger degree. The weight represents the closeness of neighbor nodes, so the local similarity should be relevant with the weight and its distribution. Based on these ideas the concept of the edge weight strength as to node pairs is proposed to measure the local similarity of neighbor nodes.

When measuring the effect of the path on the node similarity, we think the farther the two nodes are, the less the possibility of a link is between them. That is to say, the shorter path between nodes contributes more to the similarity than the longer path. Based on this idea the concept of path similarity contribution is proposed to describe the paths similarity contribution to the nodes connected by it. Moreover, the similarity of transmission nodes of multiple paths is defined based on the concept of path length to describe the total contribution of all these paths of different step-size to the similarity of node pairs.

The algorithm takes the similarity of transmission nodes of multiple paths as the link prediction score of node pairs. Then it calculates the similarity of all node pairs which have yet not established links according to the definition of similarity formula. Through sorting these scores in descending order, the node pairs with higher similarity are most likely to establish links.

4. **Implementation of the Algorithm.**

4.1. **Preliminaries.** In order to accurately describe the definition of node similarity we give the following instruction: The weighted social network is represented as G= (V, E, W) where V is the set of nodes, E is the set of edges and W is the set of edges' weights. w(v_i,v_j) represents the weight of the edge connecting the node pairs (v_i,v_j). For unweighted networks all edge weights are 1.

Definition 4.1. *Node Strength: Given a weighted social graph G=(V,E,W), $v_i \in V$, $e(v_i) \subseteq E$ is the set of edges connecting to the node v_i. The node strength of v_i is defined as:*

$$s(v_i) = \sum w(e(v_i)) \tag{1}$$

Definition 4.2. *Edge Weight Strength: Given G=(V,E,W), $v_i, v_j \in V$, the edge weight strength of node pairs (v_i, v_j) is defined as:*

$$sw(v_i, v_j) = \frac{w(v_i, v_j)}{s(v_i) + s(v_j) - w(v_i, v_j)} \tag{2}$$

In this paper, we use the edge weigh strength of node pairs (v_i, v_j) to represent the local similarity score of the two nodes, and it is denoted as lsim(v_i, v_j), namely, lsim(v_i, v_j) =

$sw(v_i, v_j)$. The bigger the edge weigh strength of node pairs is, the higher their local similarity is.

Definition 4.3. *Path Similarity Contribution: Given G=(V,E,W), $v_i, v_j \in V$. the k path connecting v_i and v_j is represented as $l_k(v_i, v_j)$=$v_i e_{ik} v_{k1} e_{k1} v_{k2} \cdots e_{kn} v_j$. The contribution of the path l_k to the similarity of node pairs (v_i, v_j) is defined as:*

$$SL_k(v_i, v_j) = lsim(v_i, v_{k1}) * lsim(v_{k1}, v_{k2}) * \cdots * lsim(v_{kn}, v_j) \tag{3}$$

Definition 4.4. *Similarity Contribution: Given G=(V,E,W), $v_i, v_j \in V$. L=$\{l_1, l_2, \cdots, l_p\}$ is the set of paths connecting v_i and v_j. The contribution of all paths connecting v_i and v_j to the similarity of node pairs (v_i, v_j) is defined as:*

$$STNMP(v_i, v_j) = \sum_{k=1}^{p} SL_k \tag{4}$$

It is called similarity contribution for short. In this paper, we use similarity contribution STNMP(v_i, v_j) as the total similarity score of node pairs (v_i, v_j) based on transmission nodes of multiple paths. The larger value of STNMP(v_i, v_j) indicates more similarity between v_i and v_j and more possibility to establish a link.

Definition 4.5. *Step Length of the Path: Given G=(V,E,W), $v_i, v_j \in V$. $l_k(v_i, v_j) = v_i e_{ik} v_{k1} e_{k1} v_{k2} \cdots e_{kn} v_j$ is a path connecting v_i and v_j. The step of the path l_k is defined as the number of edges passed through by l_k and it is denoted as $|l_k(v_i, v_j)|$.*

4.2. **Implementation Steps of the Algorithm.** The description of the algorithm are as follows:

Input: The undirected weighted network graph G and its adjacency matrix A. If the node v_i and v_j are neighbors $A_{ij} = w(v_i, v_j)$, or else A_{ij}=0.

Output: Top K node pairs that are most likely to establish links.

Step 1: Calculate the local similarity scores of all adjacent node pairs of G based on the definition 4.2 of edge weight strength in equation (2) and use the list to store the calculation results $(v_i, v_j, lsim(v_i, v_j))$.

Step 2: $\forall v_i, v_j \in V$ and $e(v_i, v_j) \notin E$, calculate all the similarity contributions for node pairs (v_i, v_j) on the condition of $|l(v_i, v_j)| \leq 6$ according to equation (4) in definition 4.4 and use the list to store the calculation results $(v_i, v_j, STNMP(v_i, v_j))$.

Step 3: Sort the values of STNMP(v_i, v_j) in descending order and take the top K node pairs as link prediction results based on transmission nodes similarity of multiple paths.

5. **Experiments and Analysis.** Based on several real datasets obtained from networks, the comparative analysis were done on the accuracy and efficiency of the STNMP algorithm with some similarity index such as CN, Jaccard, AA, the shortest path and FriendTNS algorithms using Precision and AUC as the evaluation index. The results show its higher prediction accuracy than these existing algorithms.

5.1. **Datasets.** In our experiments we used seven typical real datasets which represented different types of networks. The first four are weighted networks and the other three are unweighted networks. The network topology information of these datasets were shown in table 1.

TABLE 1. Topology information of datasets

Dataset	$\lvert V\rvert$	$\lvert E\rvert$	Network Diameter	Graph Density	Weighted Average Degree	Average Degree	Average Clustering Coefficient	Average Length of Shorest Path
Zachary′s Karate club	34	78	5	0.139	13.588	4.588	0.588	2.408
Train Bombing	64	243	6	0.121	7.594	8.812	0.711	2.691
Net Science	379	914	17	0.013	2.583	4.823	0.798	6.042
US Airports	332	2126	6	0.039	0.924	12.807	0.749	2.738
(unweighted) Zacharys Karate club	34	77	5	0.137	4.529	4.529	0.574	2.426
Dolphins	62	159	6	0.084	5.129	5.129	0.29	3.111
American College Football	115	613	4	0.094	10.661	10.661	0.403	2.508

5.2. **Division of training set and testing set.** In order to evaluate the prediction accuracy of the algorithm, we need to divide the known link set E into a training set and a testing set. The commonly used method is to divide the dataset into 10 subsets, one of which is selected as the testing set and the remaining 9 are combined together as the training set in each experiment [21]. Repeat the division mentioned above 10 times to ensure that each subset can be used as a testing set only once and all the sample data can not only be trained but also be tested.

In our experiments, for each dataset we randomly selected 10% links from the set E as the testing set and denoted it as E^{Te}. The remaining 90% links were used as the training set and we denoted it as E^{Tr}. The division should ensure that E=$E^{Te} \cup E^{Tr}$ and $E^{Te} \cap E^{Tr} = \varnothing$. Moreover, it should also ensure that the remaining 90% links which would be used as training set can form a connected graph. The links in the training set were considered as known information while the testing set was used to test and verify the prediction accuracy of the algorithm.

5.3. **Evaluation index.** There are three commonly used index for the evaluation of link prediction accuracy, namely, AUC, Precision and Ranking Score [3]. The focuses of these three index are different. The AUC index can measure the overall prediction accuracy of the algorithm. The Precision index only evaluate the prediction accuracy of the top L links and its value is largely affected by the set value of the parameter L in experiments. The Ranking Score focuses on the evaluation of the rank of these predicted links.

In our experiments, we used AUC and Precision as evaluation index simultaneously. The value of the parameter n in AUC index was set to 20000 and the value of the parameter L in Precision index was set to the half of the number of links in set E.

5.4. **Performance comparison and analysis.**
(1) choice of the steplength

In the initial experiments, we divided each dataset into training set and testing set randomly and calculated the similarity of node pairs based on transmission nodes on multiple paths of different steplength. After repeating the above steps for each dataset under the same experimental environment, we got the prediction accuracy based on AUC evaluation index and running time of the STNMP algorithm. The results were shown in figure 1, figure 2 and table 2. The data in these two graphs and table 2 were the average value of these 10 independent experimental results.

From these data we know that along with the increase of the steplength the running time of the algorithm sharply increased while the prediction accuracy decreased. For the 7 datasets used in our experiments, the prediction accuracy of the algorithm reached

TABLE 2. Running time of STNMP algorithm based on different steplength (ms)

Dataset	$\lvert 1\rvert = 2\sim 3$	$\lvert 1\rvert = 2\sim 4$	$\lvert 1\rvert = 2\sim 5$	$\lvert 1\rvert = 2\sim 6$
Zachary′s Karate club	70	324	1381	4854
Train Bombing	1431	17026	187548	1918459
Net Science	13351	90860	604354	4007165
US Airports	920590	16790903	507721016	1209501342
Zachary′s Karate club (unweighted)	92	415	1665	6846
Dolphins	360	1923	10604	52117
American College Football	6554	66929	234634	2098624

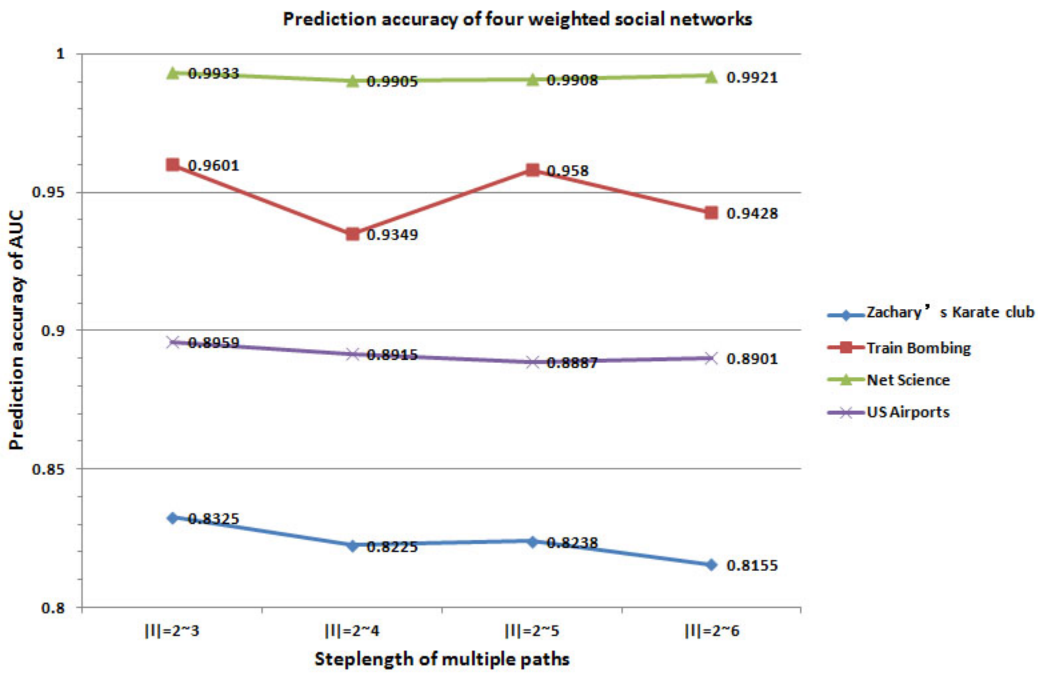

FIGURE 1. Prediction accuracy for datasets of weighted social networks

the highest value when we chose paths of $|l|$=2 and $|l|$=3 to calculate the similarity contribution. We know that the average length of the shortest path of vast majority of networks is around 3. So for the small scale networks we improved the STNMP algorithm in the our later experiments in order to avoid the high complexity caused by the calculation of path similarity contribution of all steplength. We set the upper limit of the steplength to 3. That is to say we finally defined the similarity based on the transmission nodes of multiple paths as the total contribution of those paths with $|l|$=2 and $|l|$=3 to achieve higher accuracy and efficiency.

(2) prediction accuracy

For weighted networks, we compared STNMP algorithm with three weighted similarity index of CN, AA and Jaccard [24]. For unweighted networks, we compared the algorithm with CN, AA, Jaccard, the shortest path algorithm [21] and FriendTNS algorithm [20]. With regard to each dataset the prediction accuracy based on AUC and Precision index

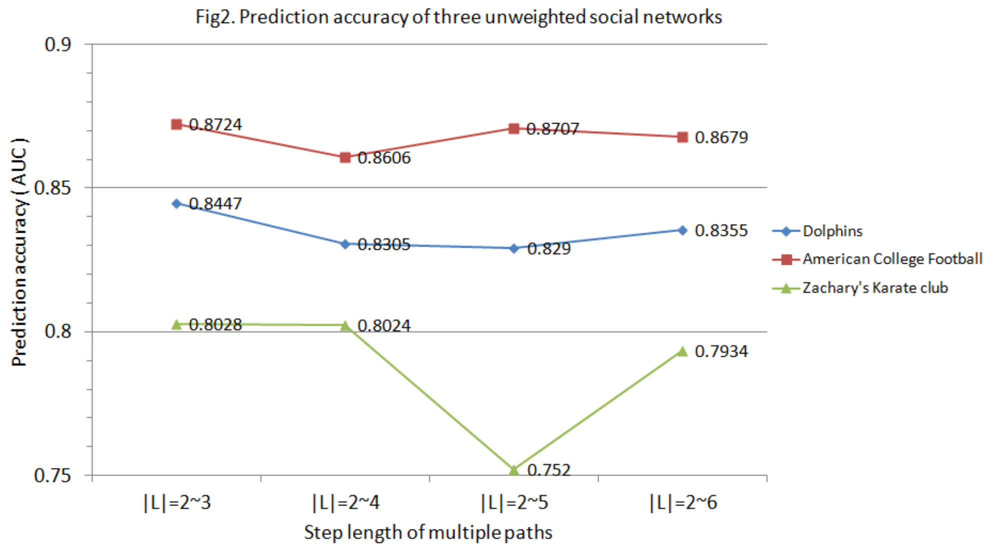

FIGURE 2. Prediction accuracy for datasets of unweighted social networks

of these algorithms were shown in table3, table4, table5 and table 6 respectively. For quick attention, the highest value in each line of these tables were shown in bold font.

TABLE 3. Prediction accuracy based on AUC (weighted networks)

Dataset	CN	Jaccard	AA	STNMP
Zachary's Karate club	0.7561	0.6351	0.7663	**0.8325**
Train Bombing	0.9207	0.9247	0.9347	**0.9601**
Net Science	0.9746	0.9743	0.9804	**0.9933**
US Airport	0.8386	0.8716	0.8803	**0.8959**

TABLE 4. Prediction accuracy based on Precision (weighted networks)

Dataset	CN	Jaccard	AA	STNMP
Zachary's Karate club	**0.2**	0	0.175	0.075
Train Bombing	0.9667	0.7583	**0.9833**	0.5417
Net Science	0.2804	0.3435	**0.5913**	0.3939
US Airport	0.1745	0.0604	0.2066	**0.2660**

TABLE 5. Prediction accuracy based on AUC (unweighted networks)

Dataset	CN	Jaccard	AA	ShortestPath	FriendTNS	STNMP
Zachary's Karate club	0.7105	0.6303	.7322	0.2992	0.7328	**0.8028**
Dolphins	0.7726	0.7704	0.7788	0.2209	0.7959	**0.8447**
American College Football	0.8368	0.8472	0.8380	0.2577	0.8194	**0.8724**

From the data we know, for these 7 networks the STNMP algorithms prediction accuracy based on AUC evaluation index is always the highest. It means that in this kind

Table 6. Prediction accuracy based on Precision (unweighted networks)

Dataset	CN	Jaccard	AA	ShortestPath	FriendTNS	STNMP
Zachary′s Karate club	**0.200**	0	0.1250	0	0.1250	0.1000
Dolphins	**0.0875**	0.050	0.0874	0	0.0375	0.050
American College Football	0.3433	**0.4333**	0.3367	0.0067	0.0733	0.4267

of network whose clustering coefficiency is relatively high and distribution of degree and weight are relatively even, the definition of the local similarity based on edge weight strength achieves a better prediction results. With regard to Precision evaluation index, because its value is greatly influenced by the value of the parameter L, the prediction accuracy of all these algorithms are all low. Compared with other algorithms, the prediction accuracy of STNMP algorithm is relatively low, but the difference is just a little. And in the US Airports network the accuracy of STNMP algorithm is the highest. In view of the fact that AUC is currently the most widely used evaluation index for the link prediction algorithms, the above experimental results about AUC data showed that the prediction accuracy of STNMP algorithm is better than these existing algorithms. It can achieve ideal prediction results in weighted and unweighted networks.

(3) complexity analysis

The STNMP algorithm used matrix and list to store information of edges and nodes and the similarity scores. Moreover, it calculated the similarity contribution of multiple paths of $|l|$=2 and $|l|$=3 for all node pairs which have not yet established links. So compared with other algorithms, the computational complexity of the STNMP algorithm increased. But for small scale networks, it can also guarantee the feasibility and effectiveness of the running time on the premise of achieving higher prediction accuracy meanwhile. Under the same experimental environment, we obtained the running time of the algorithm respectively for these datasets as shown in table 7 and table 8.

Table 7. Running time of the algorithm (weighted networks)

Dataset	CN	Jaccard	AA	STNMP
Zachary′s Karate club	1	3	1	64
Train Bombing	1	12	11	1031
Net Science	212	603	456	13351
US Airport	202	537	499	920590

Table 8. Running time of the algorithm (unweighted networks)

Dataset	CN	Jaccard	AA	ShortestPath	FriendTNS	STNMP
Zachary′s Karate club	9	3	6	1	15	92
Dolphins	31	7	7	12	53	360
American College Football	107	39	42	31	268	6554

6. **Conclusion.** A new link prediction algorithm STNMP is put forward. Firstly, the edge weight strength and path similarity contribution are defined. On basis of these concepts, the similarity of transmission nodes of multiple paths with the steplength of 2 and 3 is defined for the link prediction in weighted social networks. Experiments were

carried out on several real datasets to verify the accuracy and effectiveness of the algorithm using AUC and Precision as evaluation index. We also compare the performance of this algorithm with many classical algorithms and the results show the good generality and the higher prediction accuracy of this algorithm which needs to be furtherly improved for large scale networks. The improvement of the definition of similarity index, its application into the signed networks and analysis of the influence of the new links on overall balance of the network structure will be covered in our future research.

Acknowledgment. Project supported by the National Natural Science Fund Project ″The research on community discovery and information dissemination mechanism in online social networks based on the theme focused model″(No.61472340) and Qinhuangdao science and technology support project ″Link prediction in signed networks″(No.201502A003).

REFERENCES

[1] L. Zhang, Community detection of the signed network via dynamics[D], *Xi'an Electronic and Science University*, vol. 3, pp. 1-2, 2014,.
[2] M. W. Lan, C. P. Li, S. Q. Wang, Review of research on positive and negative links prediction algorithm in signed social networks, *Journal of Computer Research and Development*, vol. 52, no. 2, pp. 410-422, 2015.
[3] L. Y. LU, Link prediction of complex networks, *Journal of University of Electronic Science and Technology of China*, vol. 39, no. 5, pp. 651-661, 2010.
[4] Q. Yang, Study on the link prediction of social networks based on the improvement of similarity, *South China University of Technology*, vol. 6, pp. 1-5, 2015.
[5] R. R. Sarukkai. Link prediction and path analysis using markov chains, *Computer Networks*, vol. 33, pp. 377-386, 2000.
[6] J. Zhu, J. Hong, J.G. Hughes, Using markov chains for link prediction in adaptive web sites, *Lect. Notes Comput. Sci*, pp.60-73, 2002.
[7] A. Popescul, L. Ungar, Statistical relational learning for link prediction, *Proceedings of the Workshop on Learning Statistical Models from Relational Data*, New York:ACM Press, pp. 81-87, 2003.
[8] A. Clauset, C. Moore, M. E. Newman, *Journal Hierarchical structure and the prediction of missing links in networks* Nature, no. 453, pp. 98-101, 2008.
[9] R. Guimera, M. Sales-pardo, Missing and spurious interactions and the reconstruction of complex networks, *Proc Natl Sci Acad USA*, vol. 106, no. 52, pp. 22073-22078, 2009.
[10] Y. H. Li, H. L. Xiao, D. C. Li and R. X. Li, Research of dynamic link prediction method based on link importance, *Journal of Computer Research and Development*, vol. 48, pp. 40-46, 2011.
[11] M. Newman, The structure and function of complex networks, *SIAM review*, vol. 45, no. 2, pp. 167256, 2003.
[12] G. Chowdhury, L. A. Londres, Introduction to modern information retrieval, *Facet Publishing*, pp. 194-223, 2004.
[13] L. Adamic and E. Adar, How to search a social network, *Social Networks*, vol. 27, no. 3, pp. 187-203, 2005.
[14] M. Newman, Clustering and preferential attachment in growing networks, *Physical ReviewE*, vol. 64, no. 2, pp. 025102, 2001.
[15] M. Fredman and R. Tarjan, Fibonacci heaps and their uses in improved network optimization algorithms, *Journal of the ACM*, vol. 34, pp. 596-615, 1987.
[16] L. Katz, A new status index derived from social metric analysis, Psychometrika, vol.18, no. 1, pp. 39-43, 1953.
[17] L. Y. Lv, C. H. Jin, T. Zhou, Similarity index based on local paths for link prediction of complex networks, *Phys. RevE*, vol. 80, pp. 19, 2009.
[18] G. Jeh and J. Widom, Simrank: a measure of structural-context similarity, *Proc. of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 538-543, 2002.
[19] T. Zhou, L. Y. Lu, Y. C. Zhang, Predicting Missing Links via Local Information, *The European Physical Journal. B*, vol. 71, pp. 623-630, 2009.
[20] P. Symeonidis, E. Tiakas, Y. Manolopoulos, Transitive Node Similarity for Link Prediction in Social Networks with Positive and Negative Links, *Proc. the fourth ACM conference on Recommender system*, Barcelona-Spain, ACM Press, pp. 183-190, 2010.

[21] S. L. Li, Research on Link Prediction Methods Based on the Similarity[D], *Harbin Engineering University*,2012.
[22] Y. Li, Link Prediction Based on Link Dependency[D], *Jilin University*,2013.
[23] T. Murata and S. Moriyasu, Link Prediction of Social Networks Based on Weighted Proximity Measures[J]. *ACM International Conference on Web Intelligence*, pp. 85-88, 2007.
[24] L. Y. Lu and T. Zhou, Link prediction in weighted networks:The role of weak ties[J], *EPL: Europhysics Letters*, vol. 89, pp. 18001, 2010.
[25] Q. M. Zhang, S. M. Shang, L. L, Similarity-based classification in partially labeled networks, *International Journal of Modern Physics C*, vol. 21, no. 6, pp. 813-824, 2010.
[26] S. L. Zhang and Y. Zhou, Time weighted social networks link prediction algorithm based on random walk, Computer Application and Software, vol. 31, no. 7, pp.28-30, 2014.