Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **1**, January **2018**

Partitioning Weighted Social Networks based on the Link Strength of Nodes and Communities

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

Jing-Feng Guo and Jing Chen
College of Information Science and Engineering
Yanshan University
Qinhuangdao, Hebei 066004, China

Received July, 2017; revised October, 2017

Abstract. *The algorithm CD_WLS is proposed for partitioning weighted social networks reasonably and effectively. Firstly, the link strength between nodes based on their common neighbors is defined as their weighted similarity. Then nodes are clustered fast and the initial partition of the network is achieved. Finally, closely connected communities would be merged on the basis of the definition of the link strength between communities so as to optimize the initial partition and get a more accurate division result. Experiments are carried out on many artificial and real weighted networks using the weighted modularity as the evaluation criterion to verify the effectiveness and correctness of the algorithm proposed. Results show that the similarity index of weighted link strength defined in the paper is superior to WCN, WAA and WRA. Meanwhile, the speed and accuracy of CD_WLS algorithm are improved greatly compared with the WGN algorithm. Furthermore, the algorithm proposed can achieve higher accuracy for community partition in weighted networks than Lu_Algorithm and CRMA algorithm.*
Keywords: Weighted networks; Link strength; Common neighbors; Similarity; Community partition

1. **Introduction.** With the rapid development of social network applications such as Facebook and Twitter, the research on social networks has become the focus of our attention. The community structure of the social network has theoretical and practical value in understanding the topology, functional attributes and behaviors of the real network. Most existing community partition algorithms aimed at unweighted networks. However, networks in real world often have weight attributes. For example, the weight in the air transport network represents the number of flights between two airports during a given period of time. In weighted networks, the weight can not only express whether there is a relationship between two nodes, but also can express the closeness of the relationship. The information contained in the weight can better describe the structural characteristics and help to understand the nature of the network. It is also very important for community partition of the network. For example, the weight in the semantic network expresses the possibility of the two words belonging to the same semantic group. Therefore, in the clustering, if we ignore the weight, a significant amount of information would be lost, which may lead to a big gap between final clustering results and the actual semantic

structure. In addition, most researches have shown that the distribution of the weight has a great influence on the community structure and dynamic behaviors of the network. Therefore, for the community partition in weighted networks, weights should be taken as an important consideration for clustering.

The hierarchical clustering method is a classical community partition method. However, existing weighted similarity index such as weighted CN and weighted AA only considered the influence of the strength of common neighbors and the weight of the edge connecting the node with its neighbors on the similarity, which ignored the case of two nodes having no common neighbors. Therefore, for some networks, if more node pairs have no common neighbors, these indexes can't measure the similarity accurately, which may lead to a lower partition accuracy. In addition, the computational complexity of hierarchical clustering methods singly based on the modularity is high. So in order to improve the accuracy and reduce the complexity in partition of weighted networks, a community discovery method based on the weighted link strength was proposed, namely, CD_WLS (Community Discovery based on the Weighted Link Strength). Through the definition of weighted link strength between nodes based on their common neighbors, the clustering of nodes was achieved and the initial communities were found. Then the link strength between communities was defined and based on this the closely related communities would be merged if this can improve the network modularity so as to optimize the initial clustering results and achieve more accurate partitions. The related work was analyzed firstly. Then definitions and the description of the algorithm proposed were given. Finally, the effectiveness and correctness of the algorithm were verified through experiments.

2. **Related Work.** Now there have been some researches on community partition in weighted networks. Newman proposed WGN algorithm[1] through the weighted edge betweenness. Han[2] proposed the improved CNM algorithm by defining the incremental modularity about weights. Saha[3] proposed a method for soft community partition of weighted networks based on the concept of fuzzy clustering. Tushar[4] proposed the AGMA algorithm which divided weighted signed graphs into several communities according to the link type and the weight. Yao[5] proposed a community discovery method in weighted short message network and studied three weighted similarity indexes, namely WCN, WAA and WRA. Lu[6] proposed a community division algorithm (denoted as Lu_Algorithm) based on the similarity which constructed the similarity matrix and then merged the node that had the largest similarity with the current node so as to implement community partition in weighted networks. Lu[7] proposed a community discovery method based on the definition of community conductivity and then used it to determine whether a new community would form. Wang[8] proposed a center clustering algorithm based on the similarity. Liu[9] put forward an algorithm based on the attractiveness between communities. Yao[10] proposed a composite weighted model to represent the activity of stock networks and then realized the division of composite weighted networks. Guo[11] proposed an evolutionary community discovery method in dynamic weighted networks. Lu[12] defined the internal centrality and interaction centrality to describe the nodes, then proposed an efficient community discovery algorithm for weighted networks. Guo[13] proposed the CRMA algorithm which improved problems of AGMA algorithm and realized community partition of weighted networks through clustering, reclustering and merging. Mairisha[14] implemented community discovery of weighted graphs by incorporating weight variables into the mapping function of the modularity. Shen[15] defined a weighted filtering factor and implemented partition of weighted networks by iterative filtering operations. Chen[16] proposed a novel hybrid Bayesian model to implement overlapping community partition in weighted networks.

3. **Problem Definition.** In weighted networks, weights are given to edges that are called the edge weight. They can be divided into two types according to their meanings, namely, similarity weight and dissimilarity weight. In this paper, the similarity weight is used which means the larger the weight is, the higher the similarity between two nodes is. So the key of the algorithm is to effectively capture the topology attributes to define the similarity measurement so as to achieve node clustering and community partition. In view of the deficiency of existing weighted similarity index, information of the degree and the strength of nodes and their common neighbors, and the edge weight are all taken into consideration in the definition of similarity so as to reduce the computational complexity and improve the accuracy. We think that if two nodes are not connected directly, their weighted link strength is 0. If they are connected directly, the weighted link strength depends on the contribution of their common neighbors to the similarity. And if two nodes have no common neighbors, the link strength depends on the strength of the two nodes and their link weight.

Definition 1 Weighted graph. Let $G = (V, E, W)$ be an undirected and weighted graph with n nodes and m edges. $V(G) = \{v_i|1 \leq i \leq n\}$ is the node set of G. $E(G) = \{e_i|e_i = (v_x, v_y); 1 \leq i \leq m; v_x, v_y \in V(G)\}$ is the edge set of G. $W(G) = \{w_{e_i}|w_{e_i}$ represents the weight of e_i $\}$ is the edge weight set of G. $A = [A_{ij}] = [w_{ij}]$ is the adjacency matrix of G, where $w_{ij} \in [0, \infty)$ represents the weight of the edge connecting v_i and v_j. $C(G) = \{C_i(G)|C_i(G) \subseteq G, 1 \leq i \leq n\}$ is communities for the partition of G.

Firstly, we hold that the more common neighbors the two nodes have, the higher similarity they would have. Meanwhile, the influence of common neighbors with different activity on the link strength between two nodes is different. In weighted networks, the strengths of the two nodes with the same degree are not necessarily the same and vice versa. Therefore, it is believed that the common neighbor with the lower degree and the higher node strength would contribute more to the link strength between nodes than those of common neighbors with the higher degree and the lower strength. Based on this, the unit weight of the node is defined to measure the similarity contribution of the node to the other two nodes connected to it.

Definition 2 Unit strength of the node. For a given $G = (V, E, W), \forall v_x \in V(G)$, the unit strength of v_x is defined as follows.

$$u_x = \frac{s_x}{d_x} = \frac{\sum\limits_{v_y \in \Gamma(v_x)} w_{xy}}{d_x}$$

Here, $\Gamma(v_x) = \{v_y|(v_x, v_y) \in E(G), v_y \in V(G)\}$ represents the neighbor set of v_x, $s_x = \sum\limits_{v_y \in \Gamma(v_x)} w_{xy}$ represents the strength of v_x, and d_x represents the degree of v_x.

As shown in fig.1(a), the common neighbor v_{z2} contributes more to the similarity between v_x and v_y than v_{z1}. In other words, the larger the unit strength of the common neighbor is, the greater the contribution of the common neighbor to the similarity of the two nodes connected to it is.

Secondly, when the unit strength of two common neighbors are the same, the influence of the weight of the edge connecting the two nodes to their neighbors on the link strength would be considered. We hold that the greater ratio of the sum of weights of edges connecting the two nodes with their common neighbor to the sum of weights of all edges connecting to these two nodes is, the higher similarity these two nodes have. Based on this, the link coefficient of the common neighbor is defined to measure the contribution degree of the sum of weights of the two edges connecting to the common neighbor compared with the sum of edge weights of these two nodes.

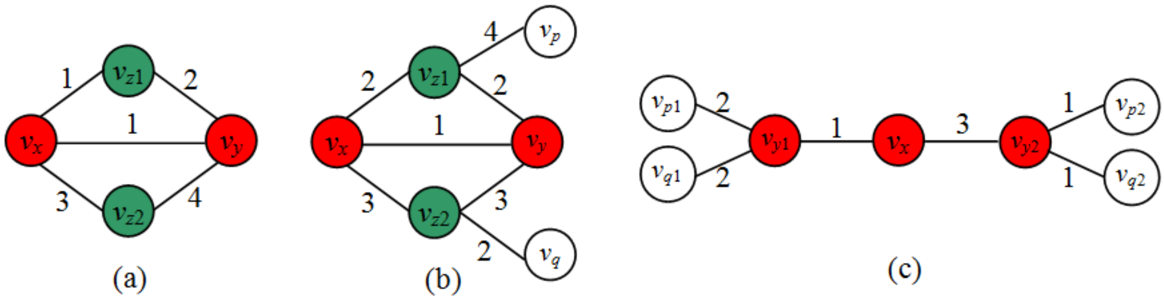

FIGURE 1. Diagrammatic sketch of the weighted link strength between nodes

Definition 3 Link coefficient of the common neighbor. For a given $G = (V, E, W)$, $v_x, v_y \in V(G)$, $\forall v_z \in V(G) \cap \Gamma(v_x) \cap \Gamma(v_y)$, the link coefficient of the common neighbor v_z to the node pair $< v_x, v_y >$ is defined as follows.

$$\varphi_{v_z}^{WCN}(v_x, v_y) = \frac{w_{xz} + w_{zy}}{s_x + s_y - w_{xy}}$$

As shown in fig.1(b), the common neighbor v_{z2} contributes much more to the similarity between v_x and v_y than v_{z1}. That is, among all common neighbors with the same unit strength, the larger the link coefficient of the common neighbor is, the greater its contribution to the similarity between the two nodes is.

Based on definition 2 and definition 3, the link strength of the common neighbor is proposed to measure the similarity contribution of the common neighbor to the two nodes connected to it.

Definition 4 Link strength of the common neighbor. For a given $G = (V, E, W)$, $v_x, v_y \in V(G)$, $\forall v_z \in V(G) \cap \Gamma(v_x) \cap \Gamma(v_y)$, the link strength of the common neighbor v_z to the node pair $< v_x, v_y >$ is defined as follows.

$$LS_{v_z}^{WCN}(v_x, v_y) = u_z \varphi_{v_z}^{WCN}(v_x, v_y)$$

Thirdly, when two nodes have no common neighbors, the influence of the weight of the edge connecting these two nodes to the similarity between the two nodes would be considered. Then the weight strength of the edge is introduced as the similarity between the two nodes.

Definition 5 Edge weight strength of the node pair. For a given $G = (V, E, W)$, $\forall v_x, v_y \in V(G)$, the edge weight strength of the node pair $< v_x, v_y >$ is defined as follows.

$$EWS(v_x, v_y) = \frac{w_{xy}}{s_x + s_y - w_{xy}}$$

When two nodes have no common neighbors, the larger the edge weight strength of the node pair is, the stronger the link closeness between the two nodes is, which means the two nodes have a higher similarity. As shown in fig.1(c), the node pairs $< v_x, v_{y1} >$ and $< v_x, v_{y2} >$ all have no common neighbors. However, based on the definition 5, the link strength between v_x and v_{y2} is greater than that of v_x and v_{y1}.

Finally, the sum of the link strength of all common neighbors of the two nodes is defined as the weighted link strength between two nodes which is finally taken as their weighted similarity.

Definition 6 Weighted link strength between nodes. For a given $G = (V, E, W)$, $\forall v_x, v_y \in V(G)$, the link strength between v_x and v_y based on their common neighbors is defined as follows.

$$WLS(v_x, v_y) = \begin{cases} 0 & (v_x, v_y) \notin E(G) \\ EWS(v_x, v_y) & (v_x, v_y) \in E(G) \wedge \Gamma(v_x) \cap \Gamma(v_y) = \Phi \\ \sum\limits_{v_z \in \Gamma(v_x) \cap \Gamma(v_y)} LS_{v_z}^{WCN}(v_x, v_y) & (v_x, v_y) \in E(G) \wedge \Gamma(v_x) \cap \Gamma(v_y) \neq \Phi \end{cases}$$

Based on above definitions, the weighted similarities between directly connected nodes of G were calculated, the node having the largest similarity with the current node would be clustered into the community, and then the initial communities formed. In the clustering process, if there are some node pairs in which the node having the largest similarity with one of the node is just another one, the two nodes in each node pair would be clustered together to form a single community. Thereout, a lot of small communities would form which may lead to a lower modularity. In view of this, the link strength between communities is defined as the similarity between these two communities, and then the tightly connected communities would be merged on condition that this operation can subsequently enhance the network modularity so as to optimize the initial clustering results. Finally, more reasonable community structures could be found.

Definition 7 Weighted link strength between communities. For a given $G = (V, E, W)$, the link strength between the community $C_p(G)$ and $C_q(G)$ is defined as follows.

$$WLS(C_p, C_q) = \frac{\sum\limits_{x=1}^{n} \sum\limits_{y=1}^{n} w_{xy}}{2 \mid V_p \mid\mid V_q \mid}$$

Here, $v_x \in C_p(G)$, $v_y \in C_q(G)$, $\mid V_p \mid$ and $\mid V_q \mid$ respectively represent the number of nodes of $C_p(G)$ and $C_q(G)$.

4. Community partition algorithm based on the weighted link strength.

4.1. CD_WLS algorithm.

Based on the definition of the link strength, the hierarchical clustering method is used to complete the clustering of nodes and community partition. The specific process is shown in algorithm 1.

Algorithm 1 CD_WLS

Input: Weighted graph $G = (V, E, W)$;

Output: Community partition results of $G, C(G) = \{C_i(G) \mid C_i(G) \subseteq G, 1 \le i \le n\}$.

1. $\forall v_x, v_y \in V(G)$, calculate $WLS(v_i, v_j)$;
2. for each $v_i \in V(G)$, set $MaxLS(v_i) = \{v_j \mid \max\limits_{v_j \in V(G)} \{WLS(v_i, v_j)\}\}$;
3. $\forall v_x \in V(G)$, set $Label(v_i) = i$;
4. $V'(G) \leftarrow V(G)$;
5. Randomly select a node $v_i \in V(G)$ and $v_i.visited = false$, set v_i as the current node, denoted it as $v_{current}$;
6. $V'(G) \leftarrow V'(G) - \{v_{current}\}$;
7. Find $MaxLS(v_{current})$ as the next node to be clustered and denoted it as v_j;
8. if $Label(v_j) \neq Label(v_{current})$ then
   $Label(v_j) \leftarrow Label(v_{current})$; $V'(G) \leftarrow V'(G) - \{v_j\}$; Set v_j as the current node $v_{current}$;
   else $V'(G) \leftarrow V'(G) - \{v_j\}$; Go to step 5;
9. Repeat step 7 and step 8 until $V'(G) = \Phi$;
10. Let $C_k(G) = \{ \underset{v_i \in V(G)}{v_i} \mid Label(v_i) = k, 1 \le k \le n\}$;
11. Calculate Q^w;
12. Take $C_p(G)$ as the current community and find $C_q(G)$ that has the largest link strength with $C_p(G)$;

13. If Q^w increases after merging $C_p(G)$ and $C_q(G)$, then $C_p(G) \leftarrow C_p(G) \cup C_q(G)$;
14. Update $C(G)$, repeat step 12 and step 13 until $C(G)$ is no longer changing.
15. Output $C(G)$.

4.2. **Computational complexity analysis.** The calculation of the algorithm mainly focuses on 3 aspects, namely, calculating the link strength between nodes, finding the next node having the largest link strength with the current node, and merging communities according to their link strength. Firstly, the link strength between any two nodes directly connected is calculated, and the time complexity is $O(m)$. Secondly, the next node in the k neighbors of the current node is searched by the linked list, and the time complexity is $O(nk)$ where $k \ll n$. Finally, the link strength between any two of p communities formed in initial partitions and the modularity after merging them are all calculated, and the time complexity is $O(p^2)$ where $p \ll n$. Therefore, the computational complexity of the algorithm is $O(m + nk + p^2)$.

5. **Experiments and analyses.** To test the performance of the algorithm proposed, experiments were carried out on several artificial and real weighted data sets. We compared WLS with the weighted similarity index of WCN, WAA and WRA. Comparative analyses were also done between the algorithm proposed and other community partition algorithms such as WGN, Lu_Algorithm and CRMA. Experimental results have showed the higher accuracy of the CD_WLS algorithm using the weighted modularity as evaluation index. The experimental hardware environment is Intel Core (TM) i5, CPU 2.3GHz and 4GB DRAM. The operating system is Microsoft Windows 7 and programming tools are MyEclipse 10, Java, Python and Gephi.

5.1. **Evaluation index.** In the research of community discovery in social networks, the modularity is a commonly used standard for evaluating quality of the community structure. Its value is always between 0.3 and 0.7. For a certain division of the network, the larger the modularity is, the more reasonable the division of the network is. That is to say a larger modularity means the relatively higher precision of community partition results. In the paper, the weighted modularity Q^w proposed by Newman[1] was used as the index for evaluating the precision of the algorithm which was defined as follows.

$$Q^w = \frac{1}{2W} \sum_{ij} (w_{ij} - \frac{w_i w_j}{2W}) \delta(C_i, C_j)$$

Here, w_{ij} represents the weight of e_{ij}, $w_i = \sum_j w_{ij}$ represents the weight of v_i which is namely s_i described earlier, $W = \sum_{ij} w_{ij}$ represents the sum of weights of all edges in $E(G)$. $\delta(C_i, C_j)$ is a function. If v_i and v_j are in the same community, $\delta(C_i, C_j)$ equals 1, otherwise it equals 0.

If all weights in the network equal 1, Q^w would be the same as Q in unweighted networks. What needs to be emphasized here is that the value of Q^w and the number of communities of division results may be different in terms of different algorithms and implementation methods. Generally speaking, the larger modularity means rather accurate number of communities and the better community structure that is closer to the real network.

5.2. **Experimental results on artificial networks.** **5.2.1. Artificial data sets.** Research has shown that in weighted networks in the real world, the degree and the strength of nodes and the weight of the edge all satisfy the power law distribution. That is to say $p(k) \propto k^{-\gamma}$, $p(s) \propto s^{-\gamma}$ and $p(w) \propto w^{-\gamma}$ where $\gamma \in [2, 3]$[21]. According to this characteristic, four kinds of networks according with the power law distribution were generated using python language and the complex network modeling tool, namely NetworkX. In terms of each type of network, ten networks were generated and the corresponding number of nodes was 100, 200, $\cdots$, 1000 respectively. Then weights were assigned to edges in these networks to form 40 artificial weighted networks, and the distributions of the weight in these four kinds of networks are as follows.

(1) The weight w_{ij} is a random integer where $w_{ij} \in [1, 10]$, hereinafter referred to as the uniform distribution network.

(2) $p(w) \propto w^{-2}$, hereinafter referred to as the power-law(2) network.

(3) $p(w) \propto w^{-2.5}$, hereinafter referred to as the power-law(2.5) network.

(4) $p(w) \propto w^{-3}$, hereinafter referred to as the power-law(3) network.

5.2.2. Effectiveness verification. Experiments were carried out on above data sets, and the Q^w and runtime were obtained as shown in fig.2 and fig.3 respectively. From which we know the algorithm achieved better partition quality on these networks. The values of Q^w are always between 0.47 and 0.63 and its average is 0.57. In addition, with the increasingly expansion of the network scale, the runtime of the algorithm increased. However, its computational complexity is far less than that of WGN algorithm which has the computational complexity of $O(mn)$. It can guarantee the feasibility and validity of time on the premise of a higher accuracy of community partition.

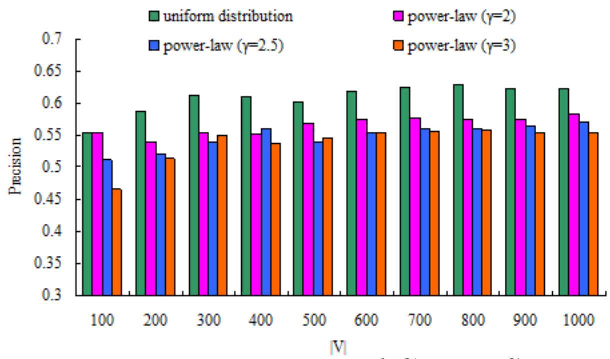

FIGURE 2. Precision of CD_WLS

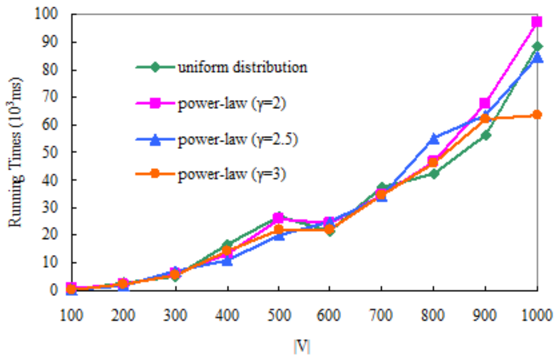

FIGURE 3. Runtime of CD_WLS

5.2.3. Accuracy comparison with other algorithms. We compared the algorithm proposed with Lu_Algorithm[7] on above artificial data sets. Both of these two algorithms have the same computational complexity order of magnitude and the results were shown in fig.4 to fig.7. From which we can see that for these 40 different weighted networks, the accuracy of the algorithm proposed is much higher than that of Lu_Algorithm.

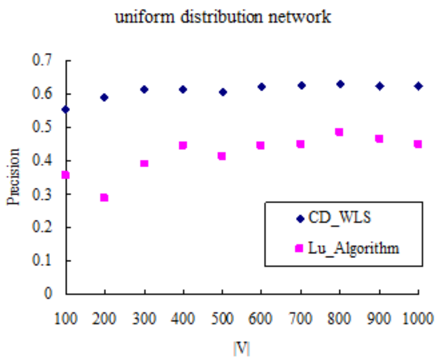

Figure 4. Precision on uniform distribution networks

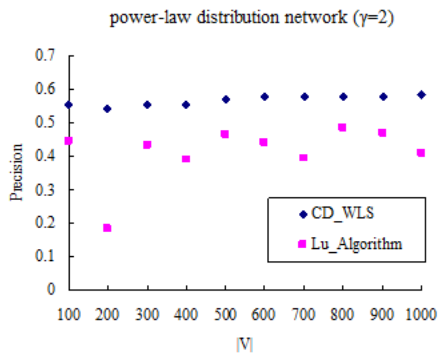

Figure 5. Precision on power-low(2) networks

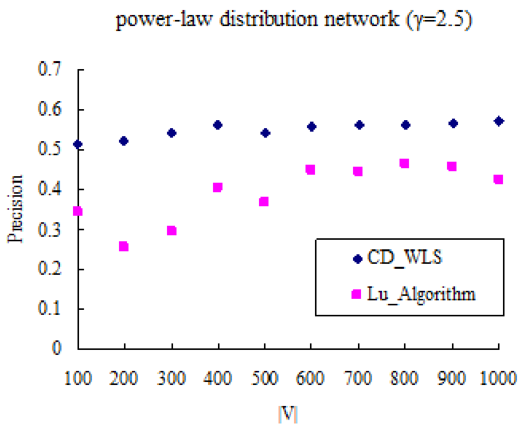

Figure 6. Precision on power-low(2.5) networks

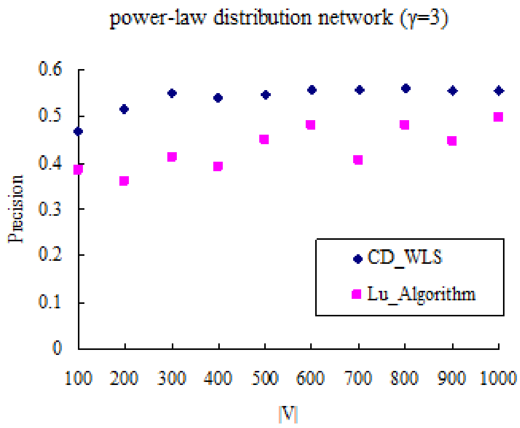

Figure 7. Precision on power-low(3) networks

5.3. **Experimental results on real networks. 5.3.1. Real data sets.** In order to further verify the correctness of the algorithm proposed, five real weighted networks were got from the Internet (http://konect.uni-koblenz.de/downloads/) for experiments. Descriptions of these data sets are as follows.

(1)Weighted Zachary's Karate club: It is a relationship network between members of a karate club in a university of America which has 34 nodes and 78 edges where a node represents a member, an edge represents the two members have close relationship and the weight represents the depth of the relationship.

(2)Train Bombing: It is a terrorist network in the train bombing in Madrid, Spain in 2004 which has 64 nodes and 243 edges where a node represents a terrorist, an edge represents two terrorists have cooperation or communication in the train bombing, and the weight represents the frequency of their contact.

(3)Les Miserables: It is a character relationship network originated from the novel of Les Miserables which has 77 nodes and 254 edges where a node represents a character, an edge represents two characters once appeared in the same scene, and the weight represents the times they appeared simultaneously.

(4)US Airport: It is a US air transport network which has 332 nodes and 2126 edges where a node represents an airport, an edge represents there is a route between two airports and the weight represents the number of flights between them.

(5)Net Science: It is a co-authored network of scientists which has 379 nodes and 914 edges where a node represents a scientist, an edge represents the two researches have ever published papers together and the weight represents the number of their cooperations.

5.3.2. Correctness verification. Community division results of WLS_CN algorithm on these real weighted networks were shown in fig.8 to fig.12 and the value of Q^w obtained were 0.4950, 0.4549, 0.5337, 0.1767 and 0.8510 respectively. In these diagrams, nodes in different communities were represented in different colors to indicate results clearly. From the experimental results, we can see that besides the US Airport network the algorithm achieved higher quality of community partition on the other four data sets. For US Airport network, the graph density is 0.039, the average degree of nodes is 12.8, and the average strength is only 0.924. In this network, 59.7% node pairs have no common neighbors. Among all node pairs with common neighbors, 46.5% node pairs have only one common neighbor, which resulted in the poor division quality of the hierarchical clustering method based on weighted similarity of common neighbors. For such networks with special topological properties, the definition of the link strength between nodes based on common neighbors needs to be further improved to achieve higher division accuracy.

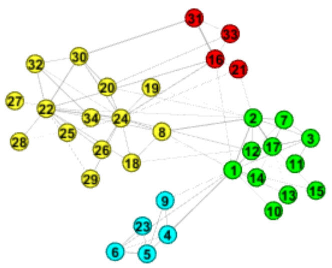

FIGURE 8. Division of club

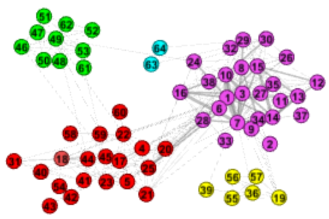

FIGURE 9. Division of Bombing

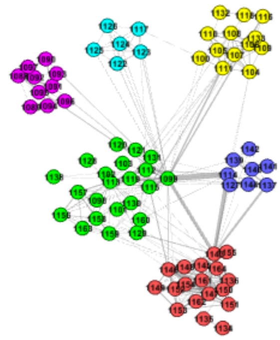

FIGURE 10. Division of Miserables

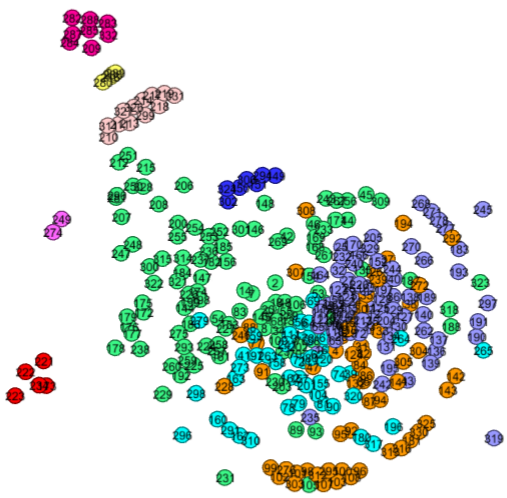

FIGURE 11. Division results for US Airport

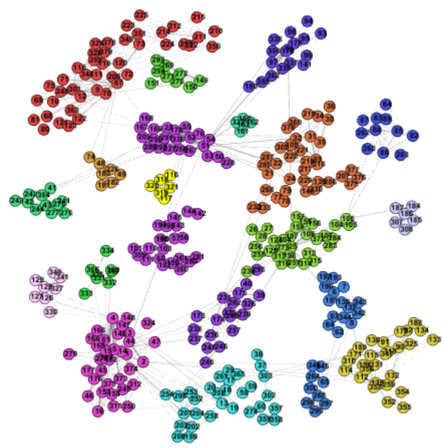

FIGURE 12. Division results for Net Science

5.3.3. Accuracy comparison with classical weighted similarity index. In view of these five networks, we compared WLS (Weighted Link Strength) defined in the paper with three weighted similarity indexes, namely WCN, WAA and WRA using the hierarchical clustering method described in this paper. Experimental results were shown in table 1 where the first column listed five data sets, the first row listed four weighted similarity indexes and the community partition results were expressed by the number of communities and the modularity of the network, namely, p/Q^w.

TABLE 1. Community partition results based on different indices

Data sets	WCN	WAA	WRA	WLS
(1) Weighted Karate Club	2 / 0.4547	2 / 0.4547	2 / 0.4547	4 / 0.4950
(2) Train Bombing	1 / 0.0303	4 / 0.3626	4 / 0.3604	5 / 0.4549
(3) Les Miserables	1 / 0.0350	3 / 0.4185	3 / 0.4577	6 / 0.5337
(4) US Airport	2 / 0.0174	3 / 0.0987	3 / 0.1039	10 / 0.1767
(5) Net Science	8 / 0.6045	19 / 0.8453	18 / 0.8499	21 / 0.8510

From table 1 we know, in Train Bombing and Les Miserables networks, the quality of partition results of WCN was the poorest because the algorithm considered only the effect of weights of edges connecting with common neighbors on the similarity. In US Airport network, many node pairs have no common neighbors and the average node strength is small, which led to the division qualities of WCN, WAA and WRA algorithm were all not ideal. However, the link strength between nodes based on common neighbors (namely WLS) defined in this paper took information of the degree of the node, the node strength and edge weights all into account. Moreover, the special case that there are no common neighbors between nodes was also processed in the algorithm. So deficiency of the former three indexes was improved and the algorithm proposed can still perform well on those special data sets. Furthermore the network modularity of community partition results of the algorithm proposed was alway the highest on these data sets.

5.3.4. Performance comparison. Aiming at these real networks, we compared the CD_WLS algorithm with the classical splitting method WGN on division accuracy and runtime. Results were shown in fig.13 and fig.14. From which we know the accuracy of CD_WLS algorithm is higher than that of WGN algorithm. With the expansion of the scale of data sets, the running time of CD_WLS algorithm increased slightly, while the running time of WGN algorithm increased significantly, which further showed that the algorithm proposed has good performance in community partition in weighted networks.

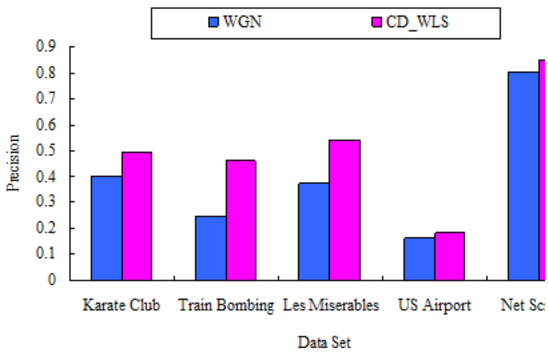

FIGURE 13. Comparison of precision

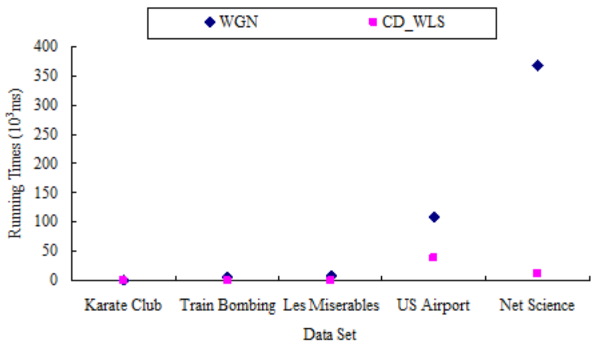

FIGURE 14. Comparison of running time

5.3.5. Accuracy Comparison. We also compared the CD_WLS algorithm with two agglomerative algorithms, namely, Lu_Algorithm and CRMA algorithm. Results were shown in fig.15 and fig.16.

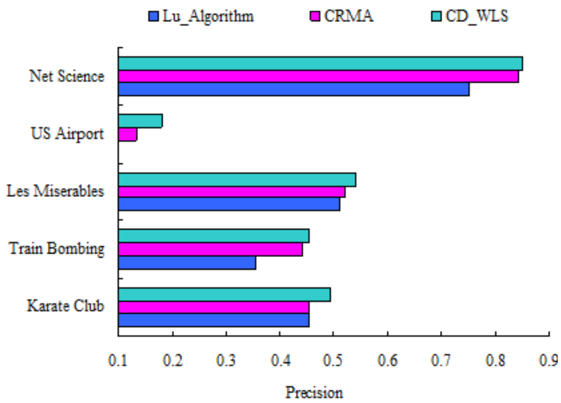

FIGURE 15. Comparison of precision

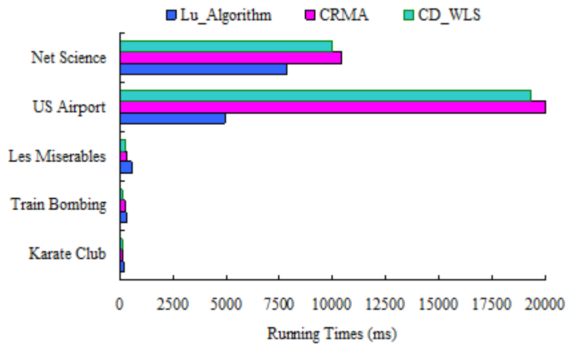

FIGURE 16. Comparison of running time

From which we know the accuracy of CD_WLS is higher than the other two algorithms. For the same dataset, the running time of CD_WLS and CRMA were slightly higher than that of Lu_Algorithm. However, these two algorithms can guarantee the feasibility and effectiveness of the time on the premise of achieving a higher accuracy.

6. **Conclusion.** The algorithm CD_WLS was proposed based on the link strength of common neighbors for community partition in weighted networks. It improved the deficiency of WCN, WAA and WRA. Comparison experiments with WGN, Lu_Algorithm and CRMA algorithm have verified the effectiveness and higher accuracy of the algorithm proposed. Though the modularity is the most common methods to measure community structure, division results with the largest modularity are not necessarily the true structure of the network such as the Zachary's club network. Therefore, using comprehensive measurements to further evaluate the performance of the algorithm is the next step.

Acknowledgment. The work was supported by the National Natural Science Fund Project (No.61472340 and No.61602401).

REFERENCES

[1] M.E.J. Newman , Analysis of weighted networks, *Physical Review E*, vol. 70, no. 5, pp. 056133, 2004.
[2] H. Han, J. Wang, H. Wang, A new algorithm to detect community structures on weighted network, *Processing of 2010 International Conference on Computational Intelligence and Software Engineering (CISE), Wuhan: IEEE*, pp. 1-4, 2010.
[3] T. Saha, C. Domeniconi, H. Rangwala, Detection of communities and bridges in weighted networks, *Springer Berlin Heidelberg*, no. 6871, pp. 584-598, 2011.
[4] T. Sharma, Lucknow, Finding Communities in Weighted Signed Social Networks, *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 978-982, 2012.
[5] Z. Q. Yao, The analysis and prediction of weighted complex networks, *Qingdao: Qingdao Technological University*, 2012.
[6] J. Lu, Y. Xu, L. P. An, A Partitioning Method for Community Structure in Weighted Networks Based on Node Similarity, *Information and Control*, vol. 41, no. 4, pp. 504-508, 2012.
[7] Z. Q. Lu, Y. G. Wen, G. H. Cao, Community detection in weighted networks: Algorithms and applicationst, *IEEE International Conference on Pervasive Computing and Communications*, vol. 26, no. 3, pp. 179-184, 2013.
[8] K. Wang, G. H. Lv, Z. W. Liang, .etc, Detecting community in weighted complex networks based on similarities, *Journal of Sichuan University (Natural Science Edition)*,vol. 51, no. 6, pp. 1170-1176, 2014.
[9] Liu Ruifang, Feng Shan, Shi Ruisheng, .etc, Weighted graph clustering for community detection of large social networks, *Procedia Computer Science*, no. 31, pp. 85-94, 2014.

[10] H. L. Yao, M. W. Luo, J. Z. Li,.etc, Hierarchical Cluster Algorithm Based on Activeness in Composite Weighted Network, *Journal of Frontiers of Computer Science and Technology*,vol. 8, no. 2, pp. 207-217, 2014.

[11] C. H. Guo, J. J. Wang, Z. Zhang, Evolutionary community structure discovery in dynamic weighted networks, *Physica A: Statistical Mechanics and its Applications*, no. 413, pp. 565-576, 2014.

[12] Z. Q. Lu, X. Sun, Y. G. Wen, .etc, Algorithms and Applications for Community Detection in Weighted Networks, *Transactions on Parallel and Distributed Systems*, vol. 26, no. 11, pp. 2916-2926, 2015.

[13] J. F. Guo, M. M. Liu, L. L. Liu, etc., An Improved Community Discovery Algorithm in Weighted Social Networks, *ICIC Express Letters*, vol. 10, no. 1, pp. 35-41, 2016.

[14] M. Mairisha, GA P. Saptawati, Improved Modularity for Community Detection Analysis in Weighted Graph, *Proceedings of the Fourth International Conference on Information and Communication Technologies (ICoICT)*, pp. 1-5, 2016.

[15] Y. Shen, Y. Liu, W. Q. Xing, Community Detection in Weighted Networks via Recursive Edge-Filtration, *Journal of Communications*, vol. 11, no. 5, pp. 484-490, 2016.

[16] Y. Chen, X. L. Wang, X. Xiang, et al, Overlapping community detection in weighted networks via a Bayesian approach, *Physica A: Statistical Mechanics and its Applications*, no. 468, pp. 790-801, 2017.

[17] D. Wang, X. Z. Jin, Research on weighted scale-free network modeling and congestion problem with adjustable clustering coefficients, *Acta phys.Sin*, vol. 61, no. 22, pp. 1-9, 2012.