ACO Based Core-Attachment Method to Detect Protein Complexes in Dynamic PPI Networks

Jing Liang[1], Xiujuan Lei[1(✉)], Ling Guo[2], and Ying Tan[3]

[1] School of Computer Science, Shaanxi Normal University, Xi'an 710119, China
xjlei@snnu.edu.cn
[2] College of Life Science, Shaanxi Normal University, Xi'an 710119, China
[3] School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China

Abstract. Proteins complexes accomplish biological functions such as transcription of DNA and translation of mRNA. Detecting protein complexes correctly and efficiently is becoming a challenging task. This paper presents a novel algorithm, core-attachment based on ant colony optimization (CA-ACO), which detects complexes in three stages. Firstly, initialize the similarity matrix. Secondly, complexes are predicted by clustering in the dynamic PPI networks. In the step, the clustering coefficient of every node is also computed. A node whose clustering coefficient is greater than the threshold is added to the core protein set. Then we mark every neighbor node of core proteins with unique core label during picking and dropping. Thirdly, filtering processes are carried out to obtain the final complex set. Experimental results show that CA-ACO algorithm had great superiority in *precision*, *recall* and *f-measure* compared with the state-of-the-art methods such as ClusterONE, DPClus, MCODE and so on.

Keywords: PPI networks · Protein complexes · Core-attachment
Ant colony optimization

1 Introduction

Protein complex is a basic structural unit that can cooperate with each other to complete the specific biological function [1, 2]. The protein-protein interaction (PPI) network [3] is composed of a number of complexes which are related to each other to perform certain functions.

In recent years, many methods have been proposed to predict protein complexes. These methods greatly promote the progress in the field of complexes prediction. According to the category of complex recognition algorithms, these algorithms can be divided into the following categories: recognition algorithm based on dense sub-graph, recognition algorithm based on hierarchical clustering, recognition algorithm based on core-attachment structure.

Based on the theory of dense sub-graph in PPI networks, lots of algorithms are proposed. In 2003, Spirin *et al*. [2] proposed the algorithm using the results of

Y. Tan et al. (Eds.): ICSI 2018, LNCS 10941, pp. 101–112, 2018.
https://doi.org/10.1007/978-3-319-93815-8_11

traversing the fully connected graphs to identify complexes. Bader *et al*. [4] proposed a method, called molecular complex detection (MCODE). Palla *et al*. [5] proposed clique percolation method (CPM) based on the close connection of sub-graph filtering algorithm. And the algorithm's application software is developed, called CFinder [6]. In 2006, Altaf-UI-Amin *et al*. [7] proposed the DPClus which can get the overlapping complexes. In 2008, Li *et al*. [8] proposed IPCA algorithm, which is based on dense sub-graph to identify overlapping complexes. In 2009, Liu *et al*. [9] proposed clustering based on maximal cliques (CMC) algorithm, which can dig out the dense sub-graph in PPI networks.

Hierarchical clustering theory is also used to predict protein complexes. In 2002, Aivan and Newman proposed GN algorithm [10], which is used to partition the modules in complex networks. In 2004, Hartuv *et al*. [11] proposed highly connected sub-graph (HCS) algorithm. In 2009, Li *et al*. [12] proposed a fast hierarchical clustering algorithm based on the local variable and edge clustering coefficient, and redefined the protein complex. In 2012, Wang *et al*. [13] proposed OMIM algorithm to predict duplicate complexes in hierarchical clustering system.

The core-attachment structure of complex is a significant view to detect protein complexes. Leung *et al*. [14] designed CORE algorithm which calculates the *p-value* for all pairs of proteins to detect cores. Wu *et al*. [15] proposed COACH algorithm.

Recently, swarm intelligence algorithms have been successfully applied to the detection of complexes in PPI networks [20]. In addition, there are many other algorithms, such as Markov Clustering (MCL) [16, 17] algorithm, ClusterONE [18] algorithm, SPICi [19] algorithm and so on.

In this paper, we proposed a protein complex prediction algorithm based on core-attachment structure and ant colony optimization method, CA-ACO. First, we adopt the weighted matrix of the dynamic PPI network as the similarity matrix of the undirected graph. Second, we use the clustering coefficient value of every node to obtain the core proteins. We mark every neighbor node of core protein with unique label through picking and dropping principle of ACO. Third, filtering processes are carried out to obtain the clustering result.

2 Methods

2.1 The ACO Based Core-Attachment Design

Since the protein is not always active in the cell cycle, in order to construct a dynamic model, we integrate the static PPI data and gene expression data because gene expression level and protein expression level are consistent. If the gene expression data at a certain timestamp is better than a threshold, then it can be considered that the protein is active at this timestamp. By using three-sigma [29] principle, active threshold is set. At a certain timestamp, if two proteins are active and interactional, it can be considered that there is an edge between the two proteins at this time. As gene expression data has 12 timestamps, the static network is divided into 12 sub-graphs which correspond to 12 timestamps. Eventually, the dynamic PPI network is constructed. Figure 1 shows a process of dynamic PPI networks construction.

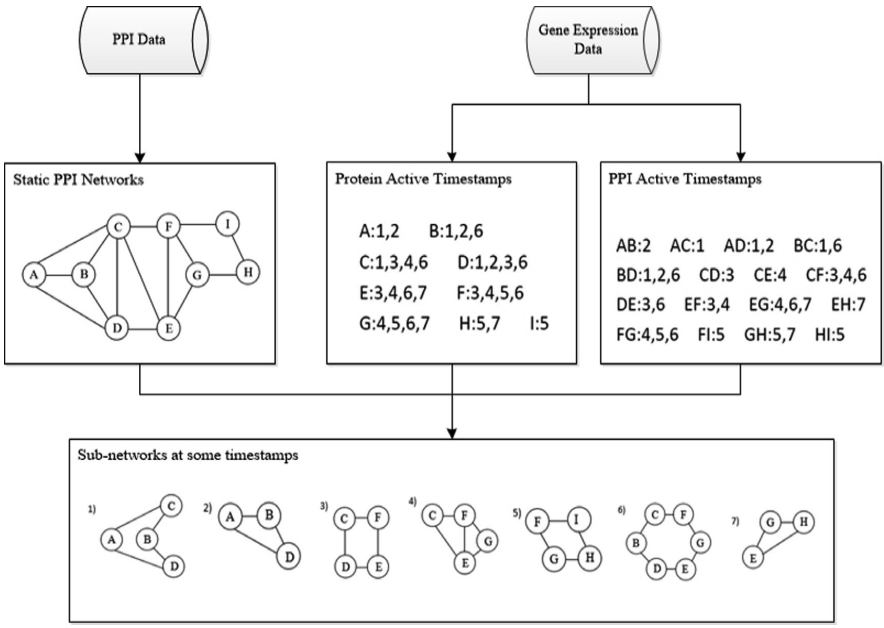

Fig. 1. An illustration example of dynamic PPI networks construction

Considering the organization of complexes, we combine the structure of core-attachment [14] with the principle of picking and dropping to predict complexes, and propose a novel algorithm, named CA-ACO. Figure 2 shows a part of the core-attachment formation design. There are two clusters which are in dotted circles of green and blue. Protein p and protein q are seed proteins. The connection between a protein and others is represented by a solid line or dotted line. The full line represents that two proteins belong to a cluster. Otherwise, they don't belong to a cluster. Taking p's neighbor protein a as an example, protein a does not belong to the cluster whose seed protein is protein p. Then, we should pick up a, and decide if protein a belongs to other cluster. Protein b is the neighbor protein of protein p and q. As the Fig. 2 shows, the protein b belongs to two clusters whose seed proteins are protein p and protein q. For protein c which is connected with protein p and protein q. Firstly, node c did not belong to cluster p and was picked up. Then, we have to decide its relationship with other seed nodes. There is a full line between c and q. Then, node c needs to be dropped out and put in the cluster whose seed node is protein q.

2.2 Description of CA-ACO Algorithm

The process of CA-ACO algorithm can be divided into 3 steps: similarity matrix initialization, clustering and purification.

In the first step, the similarity matrix of the dynamic PPI network is composed of 12 sub networks' weighted matrix. The greater the weight value is, the greater the

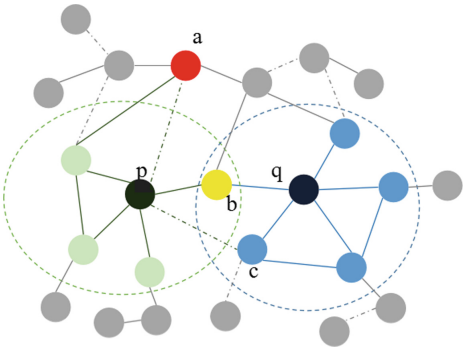

Fig. 2. A part of the core-attachment formation design (Color figure online)

similarity between the nodes is. The similarity matrix of PPI network initialization is shown in Eq. (1).

$$S(\mathrm{v}_i, \mathrm{v}_j) = \begin{cases} s(\mathrm{v}_i, \mathrm{v}_j) & if(v_i, v_j) \in E \\ 0 & else \end{cases} \tag{1}$$

Where $s(v_i, v_j)$ is weight value of edge (v_i, v_j). It represents the strength of the interaction of proteins v_i and protein v_j in weighted PPI network.

In the second step, we can obtain protein complex set from dynamic PPI network. There are two main processes: seed protein selection and attachment formation. In seed protein selection process, we need to compute the clustering coefficient value [30] of every protein by Eq. (2).

$$ccv(\mathrm{v}_i) = \frac{2 \times n_i}{|neigh(\mathrm{v}_i)| \times (|neigh(\mathrm{v}_i)| - 1)} \tag{2}$$

where $neigh(v_i)$ is neighbor nodes set of a node v_i. A protein whose clustering coefficient value is greater than the threshold will be added to the core protein set. In attachment formation step, we need to access to neighbor protein set of every core protein. By carrying out the picking and dropping operation [21] of ACO, seed proteins' neighbor proteins can be clustered. The probability of picking is calculated by Eq. (3).

$$pp(\mathrm{v}_j) = \left(\frac{k_p}{k_p + s(\mathrm{v}_i, \mathrm{v}_j)}\right)^2 \tag{3}$$

where k_p is a picking constant whose range of values is from 0 to 1, $s(v_i, v_j)$ is the similarity between the protein v_j and the current core protein v_i. In the operation of picking, the probability of picking is compared with a random probability. When the probability of picking is more than the random probability, the operation of picking is

executed. Otherwise, the protein v_j is labeled the complex whose seed protein is the protein v_i. The probability of dropping is calculated by Eq. (4).

$$pd(\mathrm{v}_j) = \begin{cases} 2 \times s(\mathrm{v}_i, \mathrm{v}_j) & if \quad s(\mathrm{v}_i, \mathrm{v}_j) < k_d \\ 1 & else \end{cases} \tag{4}$$

where k_d is a dropping constant whose range of values is 0 to 1, s(v_i, v_j) is the similarity between the protein v_j and the current core protein v_i. The probability of dropping is compared with random probability in the operation of dropping. When $pd(v_j)$ is more than a random probability, the operation of dropping is executed. Therefore, the protein v_j is labeled the complex whose seed protein is the protein v_i. Through the clustering process, we can get the initial clustering results where complexes have core-attachment structure.

In the third step, purification process is carried out. In the complex set, the complex with just one protein is deleted, and protein complex which has same proteins as others is removed. The protein complex set is obtained.

The pseudo-code of CA-ACO method is described as follows.

Algorithm: CA-ACO algorithm
Input: a dynamic PPI network *G*
Output: protein complexes set *PC*
Step 1: Initialization similarity matrix: the similarity matrix is shown in Eq.(1).
Step 2: Clustering:
    Initialization: Set various parameters.
    Seed selection: Compute the clustering coefficient of protein to obtain seed proteins by Eq.(2) .
    Attachment clustering:
        flag =0;
        For protein v_i in the core protein set
            For node v_j in the neigh(i)
                Compute the probability of picking up the protein v_j, *pp*, by Eq.(3)
                If *pp* < random value & *flag*=0
                    Then label node v_j as v_i cluster, continue
                    Else pick up node v_j , *flag* = 1
                    For other core protein v_k except protein v_i
                        Compute the probability of dropping out the protein v_j , *pd*, by Eq.(4)
                        If *pd* > random value
                            Then drop down protein v_j, label v_j as v_k , *flag*=0, continue
                        End If
                    End For
                End If
            End For
            If *flag*=1
                Drop down protein v_j, don't label and *flag*=0
            End If
        End For
        Put proteins with same labels into a cluster, get clustering results
    Return the protein complexes set *dd* from the PPI network
Step 3: Refinement: filtering the protein complexes of Step 2. Return the protein complex set *PC*

3 Experiments and Results

3.1 Experimental Dataset

In this paper, we adopt the PPI data of *S.cerevisiae* from DIP [24], MIPS [31] and Krogan database [32]. Dynamic PPI networks [3] at 12 timestamps correspond to 12 static PPI subnets. Different subnets have different size, as shown in Table 1.

Table 1. The number of proteins and interactions in each subnet of different PPI networks

Data	Timestamp	1	2	3	4	5	6	7	8	9	10	11	12
DIP	Proteins	797	941	796	623	601	530	493	944	1090	592	661	461
	Interactions	981	1444	1188	745	750	646	573	1705	2185	856	974	526
MIPS	Proteins	737	897	781	583	570	531	470	839	1014	523	616	402
	Interactions	1097	1443	1183	754	684	642	504	1238	1637	878	1207	700
Krogan	Proteins	336	379	320	256	206	189	202	580	626	304	330	250
	Interactions	334	464	331	234	210	184	213	1025	1081	314	373	258

In this paper, we use the standard known protein complex set, CYC2008 [25], which contains 408 complexes and 1,628 proteins. The biggest cluster has 81 proteins while the smallest cluster has 2 proteins in protein complexes of CYC2008.

3.2 Evaluation Criteria

The *precision* [24] indicates the proportion of the predicted protein complexes successfully matched by the standard protein complexes in the prediction of the complex. It can be defined as:

$$precision = \frac{N_{cp}}{|P|} \tag{5}$$

where $|P|$ represents the number of predicted protein complexes, and N_{cp} indicates that the number of the predicted complexes successfully matched by the known protein complexes.

The *recall* [24] indicates the proportion of the known protein complexes successfully matched by the predicted complexes in the standard of the complex. It can be defined as:

$$recall = \frac{N_{cb}}{|B|} \tag{6}$$

where $|B|$ represents the number of known protein complexes, and N_{cb} indicates that the number of the standard protein complexes successfully matched by the predicted protein complexes.

The *f-measure* [24] denotes the harmonic mean of *precision* and *recall*. It can be defined as:

$$f - measure = \frac{2 \times precision \times recall}{precision + recall} \tag{7}$$

In order to further validate the biological significance of the predicted protein complexes, we need to carry out the functional enrichment analysis by using *p-value* [36] formulated as follows:

$$p - value = \sum_{i=m}^{n} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{\mathrm{N}}{\mathrm{n}}} \tag{8}$$

where N is the number of protein in the PPI network, M is the number of proteins in a GO term, n is the number of proteins which are annotated with the same GO term. Generally, the smaller the *p-value* of a protein complex is, the stronger biological significance the complex processes will be.

3.3 Comparison with Other Methods

To evaluate the performance of CA-ACO algorithm, we compare CA-ACO with CMC [9], MCODE [4], CFinder [6], ClusterONE [18], CORE [14], COACH [15], RNSC [25], DPClus [7], MCL [16, 17], ACO-MCL [26], HC-PIN [27], MOEPGA [28] and FOCA [20] in terms of *precision*, *recall* and *f-measure* in the DIP dataset. It is obvious that the *precision* value of our method is greater than other methods *precision* value. The *recall* values of CMC, MOEPGA and FOCA algorithms are superior to our method, which are 0.5900, 0.6000 and 0.6360 respectively. However, the *f-measure* value of our method is much higher than other typical algorithms' *f-measure* value. Our method's *f-measure* value is 0.6653. It indicates that the performance of CA-ACO algorithm is optimal. The above analysis can be shown in Fig. 3.

Moreover, we also compare our method with the following prediction methods: CSO [33], ClusterONE [18], COACH [15], CMC [9], HUNTER [34] and MCODE [4] in terms of *precision*, *recall* and *f-measure* in the MIPS and Krogran dataset. As shown in Fig. 4, our method achieves the highest *f-measure* of 0.6025, *recall* of 0.5524 and *precision* of 0.6665 in MIPS dataset. On the Fig. 5, our method achieves the highest *f-measure* of 0.5844, *recall* of 0.4347 and *precision* of 0.8920 in the Krogan dataset.

We use functional enrichment analysis to validate the biological significance of methods. We calculate the *p-value* of detected complexes whose size are greater than or equal to 3. A complex is considered significant when its *p-value* is less than 0.01.

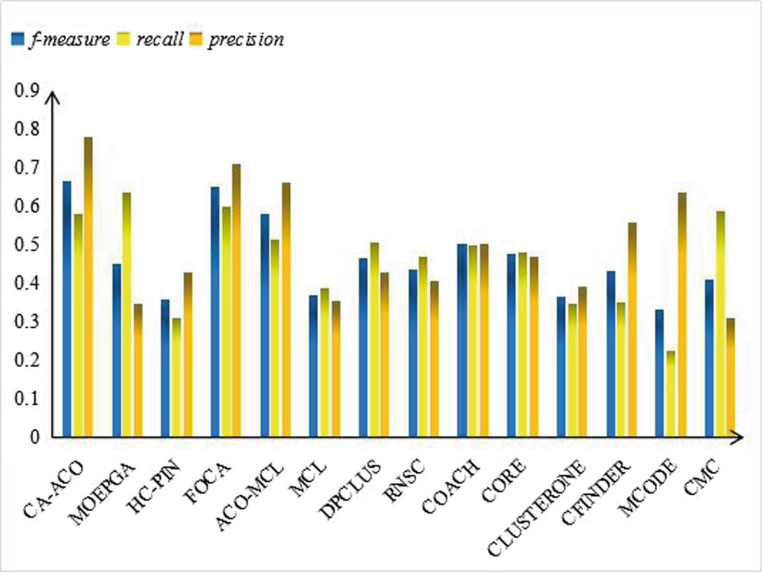

Fig. 3. *Precision*, *recall*, *f-measure* values of various algorithms on the DIP dataset

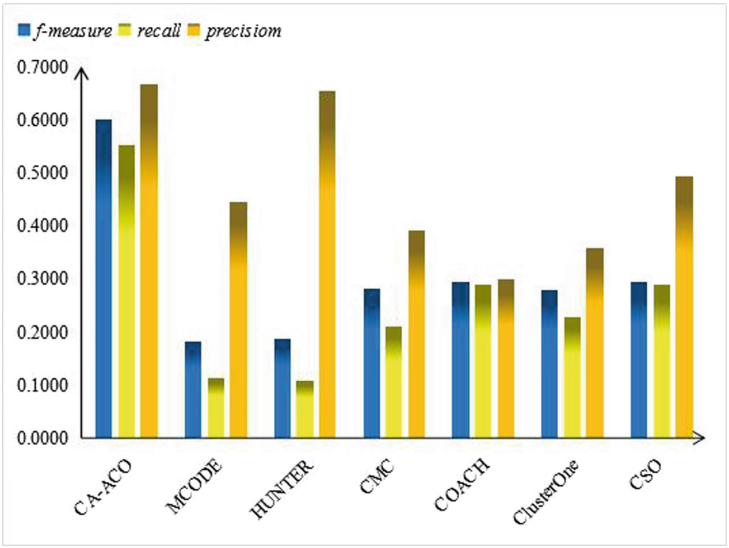

Fig. 4. *Precision*, *recall*, *f-measure* values of various algorithms on the MIPS dataset

Table 2 lists the number and percentage of the identified complexes whose *p-value* is in the range of <E−10, [E−10, E−5), [E−5, 0.01), >=0.01. Table 2 shows the comparison of the functional enrichment of complexes detected by CA-ACO, MCL, CORE and ClusterONE on DIP, MIPS and Krogan datasets. As shown in Table 2, we

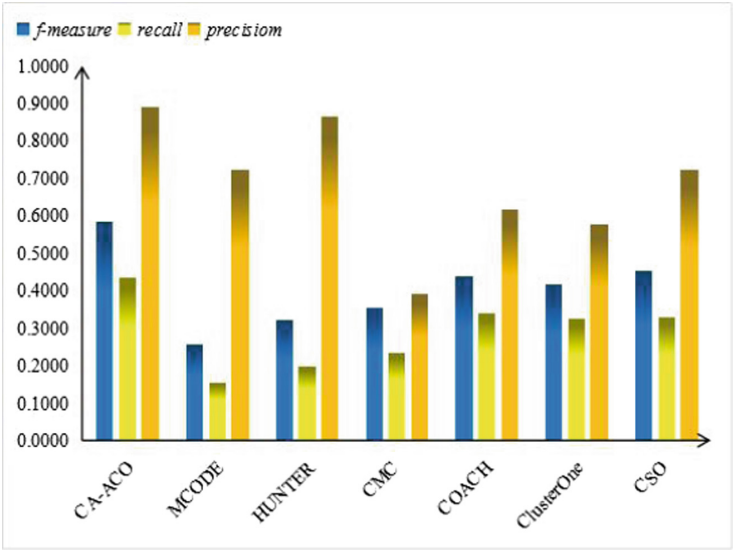

Fig. 5. *Precision*, *recall*, *f-measure* values of various algorithms on the Krogan dataset

Table 2. Functional enrichment analysis of complexes detected on DIP, MIPS and Krogan dataset

Dataset	Algorithm	PC	<E−10	[E−10,E−5)	[E−5, 0.01)	>=0.01
DIP	CA-ACO	481	15(3.12%)	107(22.25%)	254(52.81%)	105(21.83%)
	MCL	1053	66(6.26%)	183(17.38%)	362(34.38%)	442(41.98%)
	CORE	344	4(1.16%)	78(22.67%)	114(33.14%)	148(43.02%)
	ClusterONE	574	73(12.72%)	177(30.84%)	184(32.06%)	140(24.39%)
MIPS	CA-ACO	223	2(0.90%)	42(18.83%)	135(60.54%)	44(19.73%)
	MCL	606	18(2.98%)	94(15.51%)	220(36.30%)	274(45.21%)
	CORE	340	4(1.18%)	65(19.12%)	107(31.47%)	164(48.24%)
	ClusterONE	372	23(6.18%)	117(31.45%)	126(33.87%)	106(28.49%)
Krogan	CA-ACO	162	7(4.32%)	49(30.25%)	93(57.41%)	13(8.02%)
	MCL	403	59(14.64%)	103(25.56%)	119(29.53%)	122(30.27%)
	CORE	255	13(5.10%)	60(23.53%)	102(40.00%)	80(31.37%)
	ClusterONE	399	56(14.04%)	98(24.56%)	120(30.08%)	125(31.33%)

can obtain the number of predicted protein complexes by various methods on different datasets. The percentage and the amount of the predicted protein complexes with *p-value* fall into corresponding intervals. The percentage of complexes whose *p-value* is greater than 0.01 in predicted complexes by CA-ACO algorithm is the smallest. So, most of the predicted protein complexes by CA-ACO are meaningful. These illustrate that our proposed algorithm is competent to identified significant protein complexes in dynamic PPI networks.

4 Conclusions

Many of the current methods predicting protein complexes are running in a static PPI network, which ignoring the dynamic properties of the PPI network and the inherent organization of the protein complex. In this paper, we proposed a novel method for detecting protein complexes in dynamic protein interaction networks, CA-ACO, which is based on the core-attachment structure of protein complexes. We compare the performance of the CA-ACO algorithm with other state-of-the-art methods in DIP, MIPS and Krogan dataset. Experimental results show that CA-ACO algorithm is obviously superior to other methods. In addition, the shift from static PPI networks to dynamic PPI networks is important to analyze the biological significance of complexes identified from PPI networks. In the future, we will further optimize our algorithm to improve the efficiency of algorithm and the effect of biological research.

References

1. Gavin, A.C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al.: Functional organization of the yeast proteome by systematic analysis of protein complexes. Nature **415**, 141–147 (2002)
2. Spirin, V., Mirny, L.A.: Protein complexes and functional modules in molecular networks. Proc. Nat. Acad. Sci. U.S.A. **100**, 12123–12128 (2003)
3. Wang, J., Peng, X., Peng, W., Wu, F.X.: Dynamic protein interaction network construction and applications. Proteomics **14**, 338–352 (2014)
4. Bader, G.D., Hogue, C.W.V.: An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinform. **4**, 2–28 (2003)
5. Palla, G., Derényi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. Nature **435**, 814–818 (2005)
6. Adamcsek, B., Palla, G., Farkas, I.J., Vicsek, T.: CFinder: locating cliques and overlapping modules in biological networks. Bioinformatics **22**, 1021–1023 (2006)
7. Altaf-Ul-Amin, M., Shinbo, Y., Mihara, K., Kurokawa, K., Kanaya, S.: Development and implementation of an algorithm for detection of protein complexes in large interaction networks. BMC Bioinform. **7**, 207–219 (2006)
8. Li, M., Chen, J., Wang, J., Chen, G.: Modifying the DPClus algorithm for identifying protein complexes based on new topological structures. BMC Bioinform. **9**, 398–413 (2008)
9. Liu, G., Wong, L., Chua, H.N.: Complex discovery from weighted PPI networks. Bioinformatics **25**, 1891–1897 (2009)
10. Girvan, M., Newman, M.E.J.: Community structure in social and biological networks. Proc. Nat. Acad. Sci. **99**, 7821–7826 (2002)
11. Hartuv, E., Shamir, R.: A clustering algorithm based on graph connectivity. Inf. Process. Lett. **76**, 175–181 (2000)
12. Li, M., Wang, J., Chen, J., Pan, Y.: Hierarchical organization of functional modules in weighted protein interaction networks using clustering coefficient. Bioinform. Res. Appl. **5542**, 75–86 (2009)
13. Wang, X., Li, L., Cheng, Y.: An overlapping module identification method in protein-protein interaction networks. BMC Bioinform. **13**, S4 (2012)
14. Leung, H.C., Xiang, Q., Yiu, S.M., Chin, F.Y.: Predicting protein complexes from PPI data: a core-attachment approach. J. Comput. Biol. **16**, 133–144 (2009)

15. Wu, M., Li, X., Kwoh, C.K.: A core-attachment based method to detect protein complexes in PPI networks. BMC Bioinform. **10**, 169–184 (2009)
16. Leal, J.P., Enright, A., Ouzounis, C.A.: Detection of functional modules from protein interaction networks. Proteins Struct. Funct. Bioinform. **54**, 49–57 (2003)
17. Van Dongen, S.M.: Graph clustering by flow simulation (2001)
18. Nepusz, T., Yu, H., Paccanaro, A.: Detecting overlapping protein complexes in protein-protein interaction networks. Nat. Meth. **9**, 471–472 (2012)
19. Jiang, P., Singh, M.: SPICi: a fast clustering algorithm for large biological networks. Bioinform. **26**, 1105–1111 (2010)
20. Lei, X., Ding, Y., Hamido, F., Zhang, A.: Identification of dynamic protein complexes based on fruit fly optimization algorithm. Knowl. Based Syst. **105**, 270–277 (2016)
21. Leumer, E., Faieta, B.: Diversity and adaption in populations of clustering ants. In: Proceedings of the 3rd International Conference on Simulation of Adaptive Behavior: From Animal to Animals, pp. 499–508. MIT Press, Cambridge (1994)
22. Xenarios, L., Salwínski, L., Duan, X.J., Higney, P., Kim, S., Eisenberg, D.: DIP: the database of interaction proteins: a research tool for studying cellular networks of protein interactions. Nucleic Acids Res. **30**, 303–305 (2002)
23. Pu, S., Wong, J., Turner, B., Cho, E., Wodak, S.J.: Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res. **37**, 825–831 (2009)
24. Keretsu, S., Sarmah, R.: Weighted edge based clustering to identify protein complexes in protein–protein interaction networks incorporating gene expression profile. Comput. Biol. Chem. **65**, 69–79 (2016)
25. King, A.D., Pržulj, N., Jurisica, I.: Protein complex prediction via cost-based clustering. Bioinformatics **20**, 3013–3020 (2004)
26. Seçkiner, S.U., Eroglu, Y., Emrullah, M., Dereli, T.: Ant colony optimization for continuous functions by using novel pheromone updating. Appl. Math. Comput. **219**, 4163–4175 (2013)
27. Wang, J., Li, M., Chen, J., Pan, Y.: A fast hierarchical clustering algorithm for functional modules discovery in protein interaction networks. Comput. Biol. Bioinform. **8**, 607–620 (2011)
28. Cao, B., Luo, J., Liang, C., Wang, S., Song, D.: MOEPGA: a novel method to detect protein complexes in yeast protein–protein interaction networks based on MultiObjective Evolutionary Programming Genetic Algorithm. Comput. Biol. Chem. **58**, 173–181 (2015)
29. Vlasblom, J., Wodak, S.J.: Markov clustering versus affinity propagation for the partitioning of protein interaction graphs. BMC Bioinform. **10**, 99 (2009)
30. Dimitrakopoulosa, C., Theofilatosa, K., Pegkasb, A., Likothanassis, S., Mavroudi, S.: Predicting overlapping protein complexes from weighted protein interaction graphs by gradually expanding dense neighborhoods. Artif. Intell. Med. **71**, 62–69 (2016)
31. Güldener, U., Münsterkötter, M., Oesterheld, M., Pagel, P., Ruepp, A., Mewes, H., et al.: MPact: the MIPS protein interaction resource on Yeast. Nucleic Acids Res. **34**, 436–441 (2006)
32. Krogan, N., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., et al.: Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature **440**(7084), 637–643 (2006)
33. Zhang, Y., Lin, H., Yang, Z., Wang, J., Li, Y., Xu, B.: Protein complex prediction in large ontology attributed protein-protein interaction networks. IEEE/ACM Trans. Comput. Biol. Bioinform. **10**, 729–741 (2013)
34. Chin, C., Chen, S., Ho, C., Ko, M., Lin, C.: A hub-attachment based method to detect functional modules from confidence-scored protein interactions and expression profiles. BMC Bioinform. **11**, S25 (2010)

35. Wang, J., Peng, X., Li, M., Pan, Y.: Construction and application of dynamic protein interaction network based on time course gene expression data. Proteomics **13**(2), 301–312 (2013)
36. Shen, X., Yi, L., Jiang, X., Zhao, Y., Hu, X., He, T., Yang, J.: Neighbor affinity based algorithm for discovering temporal protein complex from dynamic PPI network. Methods **110**, 90–96 (2016)