

A Survey of Computational Methods for Protein Complexes Prediction Based on Static PPI Networks

Yang Yu

Software College, Shenyang Normal University, Shenyang Liaoning
Email: yuyangsd1204@126.com

Received: Jun. 6th, 2018; accepted: Jun. 20th, 2018; published: Jun. 27th, 2018

Abstract

Protein complexes are formed by interacting proteins and exhibit diverse biological functions. Protein complexes are predicted by computational methods from biological networks, which is not only important for understanding the mechanisms of biological activities and the pathogenesis of diseases, but also for making up the deficiencies of biological high-throughput experimental methods. In this paper, two types of prediction methods based on static network protein complexes are introduced and analyzed. Secondly, we discuss the deficiencies of protein complex algorithms and the challenges of this field.

Keywords

Protein-Protein Interaction Network, Clustering, Complex Prediction, Computational Methods

基于静态网络的蛋白质复合物预测方法综述

于 杨

沈阳师范大学, 软件学院, 辽宁 沈阳
Email: yuyangsd1204@126.com

收稿日期: 2018年6月6日; 录用日期: 2018年6月20日; 发布日期: 2018年6月27日

摘 要

蛋白质复合物通过相互作用蛋白质形成, 表现出多样的生物功能。使用计算方法从生物网络中预测蛋白质复合物不仅对于理解生物活动的机制和疾病的发病机理具有重要意义, 而且可以弥补生物高通量实验

方法的不足。本文介绍分析两类基于静态网络蛋白质复合物预测的方法，讨论蛋白质复合物预测算法的不足，进一步分析探讨蛋白质复合物预测所面临的挑战。

关键词

蛋白质网络，聚类，复合物预测，计算方法

Copyright © 2018 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

蛋白质复合物是含有多个相互作用蛋白质并执行不同细胞功能的生物分子[1]。从蛋白质相互作用 (PPI)网络中识别蛋白质复合物可以更好地理解蛋白质在不同细胞系统中的作用。复合物预测技术包括生物实验技术和计算技术。早期研究人员主要是通过生物实验测定蛋白复合物，但此类技术周期长，成本较高，有些复合物需要在特定的环境下才能被发现。高通量技术和计算技术的发展产生了大量的 PPI 数据，如何有效地利用这些数据构建蛋白质相互作用网络并预测复合物成为了一个巨大的挑战。计算方法预测蛋白质复合不仅可以弥补生物实验技术的不足，而且有助于研究和了解复杂的更高层次细胞组织，预测未知蛋白质功能，研究疾病的发生机理和寻找新的药物靶标。整个蛋白质相互作用网络可以看成是一个图，其中蛋白质对应于网络中的点，相互作用对应于网络中的边[2]。复合物预测是利用以机器学习、数据挖掘和复杂网络理论为基础的聚类技术从 PPI 网络数据中提取出具有相似功能蛋白质集合的过程。

近十年期间，涌现了大量的复合物预测的计算方法。本文将静态网络蛋白质复合物预测的方法分成两大类进行分析，分类如图 1 所示，并进一步对蛋白质复合物预测所面临的挑战与未来的研究方向进行探讨。

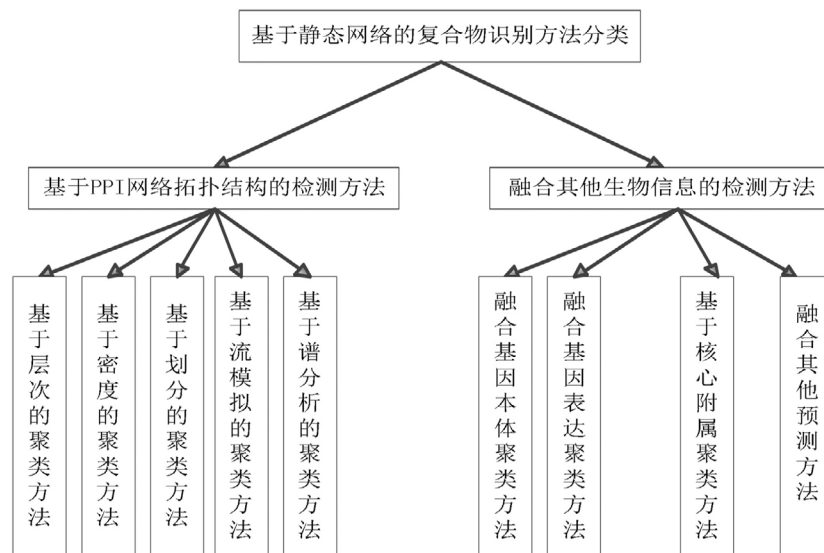


Figure 1. Classification of protein complex detection

图 1. 蛋白质复合物预测算法分类

2. 基于静态网络的蛋白质复合物预测方法

2.1. 基于网络拓扑结构的复合物预测方法

1) 基于划分的预测方法

基于划分的预测方法将网络划分成不同的簇，需要预先确定模块的数量和目标函数。当同个模块内部对象之间的相似度为最大，各模块间的相似度为最小时，则目标函数最小。Vlasblom J 等人采用近邻传播方法划分蛋白质相互作用网络[3]。Dunn 等人采用边介数的算法将蛋白质网络划分成不同的蛋白质簇[4]。

基于划分的预测方法简单且易于理解，但是不能预测重叠的蛋白质模块。因此，如何预测重叠模块仍是此类方法需要完善的问题。

2) 基于密度的预测方法

在一个簇内蛋白质倾向于表现较多相互作用，但与簇外蛋白质存在相对较弱的相互作用。在此基础上，研究人员提出相应的方法预测蛋白质复合物。

MCODE 检测网络中链接相对密集的簇为蛋白质复合物[5]。Adamcsek 等人基于团渗透方法设计在蛋白质相互作用网络中预测蛋白质复合物的 Cfnder 工具[6]。DPCLus 通过簇边界跟踪机制设计预测复合物[7]。PCP 采用基于 FS 权重的方法构造加权图，然后在加权网络中通过稠密子图预测蛋白质复合物[8]；Abdullah 等人首先获得信息富集的蛋白质，并使用聚类系数定位与蛋白质相互作用的最近蛋白质簇预测蛋白质复合物[9]；Ovrlp 基于桥蛋白概念提出计算重叠簇方法预测网络的模块化结构[10]；MINE 算法利用改进的节点权值和网络模块测量方法定义模块的边界，从而避免了邻接点的错误计算[11]；PE-WCC 首先计算基于新测量的相互作用数据的可信度，然后使用加权聚类系数的概念决定哪些子图更接近最大团[12]；王建新，李敏等人[13]通过考虑子图的密度和模块性来定义子图的适应度的 LF_PIN 算法，该算法通过将每个种子边缘延伸到一个子图，直到其适应度达到局部最大值实现复合物预测。Hernandez 等人[14]提出挖掘高度连通子图算法，该算法将 PPI 网络转换成有向无环图，减少表示的边的数目和识别子图的搜索空间，并考虑加权和未加权的 PPI 网络预测蛋白质复合物。

此类方法可以在不考虑低密度簇的情况下找到高密度的蛋白质团簇，如何找到其它拓扑类型模块是此类方法研究的重点目标。

3) 基于层次的预测方法

这种方法已广泛应用于复杂网络的研究，如社交网络和生物网络。

GN 方法不断删除网络中介数最大的边，直至删除网络中的所有边[15]。文献[16]利用 Suite Jerarca 有效地将相互作用的网络单元转化成树状图，进而通过层次聚类预测网络模块。文献[17]提出基于边的层次聚类方法预测重叠的蛋白质复合物。文献[18]考虑对共现复合物打分的方式，采用基于扩展传统集成聚类方法测蛋白质复合物。

层次聚类方法包含基于节点和基于边的聚类。虽然通过使用层次树显示一个蛋白质相互作用网络，但是当使用基于节点的层次聚类方法时，通常并不容易检测到重叠的蛋白质模块。

4) 基于流模拟的预测方法

流模拟方法模拟相互作用网络内的生物学或功能流，并尝试以形成一个模型预测各种受外部刺激产生复杂网络的行为。

马尔可夫聚类可以在有权或者无权图中模拟节点随机游动预测蛋白质复合物[19]。TRIBE-MCL 通过使用序列相似性构图计算随机行走节点概率定位蛋白质模块。受流模拟算法启发，Hwang 等人通过构建动态信号转导的模型与图的拓扑信息迭代预测蛋白质复合物[20]。CASCADE 方法通过在蛋白质相互作用

网络中计算共享概率调节转导预测蛋白质复合物[21]。Shen 等人[22]提出了一种基于迭代执行 R-MCL 算法实现软聚类, 允许识别高度重叠的簇。Ochieng 等人引入马尔可夫图聚类算法识别高度相互连接 PPI 网络中的蛋白质复合物, 算法首先构建 PPI 网络建立几何网络, 然后利用马尔可夫聚类在构建的网络上检测蛋白质复合物[23]。

流模拟方法不强调聚类内连接和节点密度, 从而避免了小规模或只有一个节点聚类结果。因为所有节点的信息流需要考虑, 该方法的时间复杂度非常高。此外, 一些蛋白质在聚类过程中会被丢掉。

5) 基于谱聚类的预测方法

谱聚类方法采用矩阵理论和线性代数理论为研究基础, 近年来已被广泛的应用到复杂网络和生物网络等研究领域。

文献[24]将谱聚类算法应用在果蝇的蛋白相互作用网络上, 发现离散的谱密度与网络的拓扑结构相对应, 擅长描述环状拓扑结构。高琳等提出构建相似性矩阵和设定聚类数目, 应用于谱聚类方法确定蛋白质模块[25]。文献[26]提出基于扩散模型的谱聚类算法, 它将蛋白相互作用网络的聚类结构解析为扩散过程中的随机游走问题。

谱聚类方法被用于计算和构造特征向量, 该方法预测的簇通常不重叠。从理论上讲, 它也能收敛在全局最优解的任何形状的样本空间, 但聚类初始化、聚类的个数、邻接矩阵构造以及特征向量的选择将直接影响最终聚类效果。因此, 如何确定这些问题和预测重叠簇仍是待解决的关键问题。

6) 基于群智能的预测方法

群智能算法是基于生物学行为的一种全局概率搜索算法, 该类算法考虑每个可用的解决方案作为一种生物实体, 解决方案的搜索和优化过程被认为是进化过程, 该方法已经在生物网络中得到有效应用。

ACOPIIN 方法结合蚁群算法与旅行商方法在蛋白相互作用网络中预测群集的节点。还有一些方法通过与功能流程聚类, 将人工蜂群算法引入到蛋白质相互作用网络预测蛋白质复合物[27]。雷秀娟等提出了一种结合萤火虫算法和基于同步的层次聚类(SHC)的复合物预测算法[28], 首先, 通过谱聚类对 PPI 数据进行预处理, 将高维相似度矩阵转换为低维矩阵。然后 SHC 算法用于执行聚类并采用萤火虫算法优化识别过程。这种方法已经有效地应用在复合物预测方面, 如何提高蛋白质检测的性能, 以及如何有效的引入生物信息仍是一个新的挑战。

2.2. 融合其他生物特征的预测方法

结合蛋白质相互作用外的其他生物学特征提高蛋白质复合物预测的准确性已成为近年来的热点研究。根据现有的相关研究, 将预测方法分为四类进行分析。

1) 融合基因表达的预测方法

基于相互作用蛋白质具有相似表达的理论, 相关研究人员融合基因表达信息预测蛋白质复合物提高预测方法的性能。

Feng 等人融合基因表达信息和蛋白相互作用数据构建有权图并设计图分割算法预测蛋白质复合物[29]。Li 等人融合基因表达谱和蛋白相互作用建立一个时间序列的网络, 设计算法预测蛋白质复合物[30]。WEC 方法基于相互作用蛋白质的边权值识别复合物, 其中权值由相互作用的蛋白质之间的边聚集系数和基因表达计算获取[31]。文献[32]使用一个新的策略来确定蛋白质复合物, 其将时间序列的基因表达数据整合到静态蛋白质相互作用数据中, 构建了一系列时间序列的子网络, 然后结合网络拓扑结构和基因本体信息确定 PPI 网络中蛋白质之间的距离, 聚类中心通常被认为是密度高于邻接点的节点, 并且与其他邻居相比到聚类中心具有较大的距离, 最后, 基于 brainstorming 方法更新初始簇, 最后优化合并簇。

2) 融合功能信息的预测方法

同一个复合物内的蛋白质通常具有相同或类似功能[33]。基于此理论,研究人员通常从两个方面融合功能信息预测蛋白质复合物:一方面是与拓扑结构相结合构建加权图;另一方面通过使用功能注释辅助预测复合物;这两类均弥补基于拓扑结构预测的不足。

NACO-FMD 方法基于蚁群算法融合拓扑和功能信息预测蛋白质复合物,并进一步提出了一种将蚁群优化与多智能体进化相结合的新算法探索和搜索空间预测功能模块。该方法具有蚁群算法和多智能体算法两者的优点,并具有良好的表现性能[34]。文献[35]基于多智能体进化提出了一种复合物预测算法。Yang 等人设计一种融合功能和拓扑结构特征捕获蛋白质相互作用网络中的功能模块,预测生物网络中具有相似功能蛋白质集合[36]。YU 等人结合蛋白质相互作用和蛋白质功相似性构建蛋白质网络,基于现实中真实的复合物信息,采用有监督的方法预测蛋白质复合物[37]。Hu 等人发现:大多数蛋白质复合物预测方法中蛋白质仅在功能信息类别的特定子集中相似,而不是整个集合,在此基础上,先定量设计每个蛋白质的偏好向量,以在确定该蛋白质所属的蛋白质复合物时定量表示每个功能类别的偏好;然后综合蛋白质的功能偏好和 PPI 网络的图形拓扑结构,设计算法 DCAFP 将蛋白质复合物预测为一个约束优化问题的问题[38]。Li 等人将蛋白质相互作用网络拓扑结构与基因本体标注信息相结合,构建了一个多关系蛋白质相互作用网络;然后设计了基于多关系蛋白相互作用网络预测具有高内聚性和低耦合性的蛋白质复合物[39]。DGGPWN [40]首先使用基于基因本体的方法测量相互作用蛋白质之间的语义相似性,并使用相似度值作为它们的权重,提出基于加权 PPI 网络的聚类算法。

3) 基于核心-附属的预测方法

Gavin 等人研究表明蛋白质复合物由核心和附属两部分组成[41]。核心部分模块内的蛋白质具有高度共表达和功能相似性,具有相对更多相互作用;而每个附属蛋白辅助核心形成一个生物学上有意义的复合物。

CORE 方法在蛋白相互作用网络中首先预测复合物的核心部分,然后预测适合于核心的附件,计算并排序检测到的簇,最后形成蛋白质复合物[42]。COACH 方法首先在邻接图中定义内核节点,并检测一些蛋白质复合物为核心蛋白质,附属蛋白质逐个连接到核心蛋白质所代表的复合物中从而形成具有生物学意义的功能结构[43]。Ma 等人首先利用蛋白质相互作用网络的特征值和特征向量构建一个虚拟网络,然后基于核心-附属关系预测复合物[44]。Kang 等人基于马尔科夫聚类设计有权图识别,利用核心-附属关系预测蛋白质复合物[45]。文献[46]提出结合基因表达数据的核心-附属预测方法。

4) 融合其他信息的预测方法

除了前面提到的三类方法外,还有融合其他的生物信息的方法。

比尔等人构造相互作用-域-相互作用网络结构预测蛋白质复合物[47]。Li 等人使用一个多层神经网络利用半监督方法预测潜在蛋白质模块[48]。Liu 等人采用一个监督图聚类预测蛋白质复合物[49]。YU 等人提出融合蛋白质氨基酸序列和拓扑信息构建有权图预测蛋白质复合物方法[50]。

综上所述,融合生物信息的方法主要用于三个方面:一构建可靠的蛋白质相互作用网络;二引入生物信息辅助复合物预测;三结合拓扑和生物信息表示蛋白复合物特征。所有这些都是用于降低噪声相互作用数据对预测性能影响,提高预测的覆盖度和准确率。

3. 复合物预测存在的问题与研究展望

随着蛋白质相互作用的网络数据的不断扩大,计算方法可以补充实验方法的局限性并进行预测复合物。近年来,图论、信息论、机器学习、数据挖掘和复杂的网络已经被引入到这一领域,并已在蛋白复合物的预测方面显示出较好的效果。但是,仍然存在一些问题,这一领域未来的挑战将需要重点关注以下几个方面:

1) 数据集不可靠性

蛋白质相互作用数据具有一定的噪声,这将会对使用计算方检测复合物产生一定的影响,因而可以从下面三个方面降低噪声数据的影响:一验证相互作用的可靠性,二设计反映不同数据源信息权重的有权图,三设计有效的预测方法。

2) 蛋白质复合物的表示

大部分聚类方法基于无监督的方法和预定义的规则通过预测稠密子图预测蛋白质复合物,这些预测到的复合物与网络中真实的复合物的并不总是一致。因此,可以通过融合多种生物信息形成合适的模型检测蛋白质复合物。

3) 融合拓扑性质和生物信息

目前有很多研究的蛋白复合物预测的方式采用了多种生物信息,如拓扑性质、功能特性和基因本体等。这些方法主要分成三类:一种是融合多元特征,这种方法主要采用半、有监督机器学习的方法从已知的复合物中得到模型函数,然后进行蛋白质复合物预测。第二类方法将蛋白质相互作用的噪声数据进行优化进而构造加权图,然后应用聚类算法预测复合物。第三类在聚类过程中引入了生物信息作为聚类条件,辅助复合物预测。因而如何有效地利用生物和拓扑特征预测复合物仍然是一个重要的研究方向。

4) 构建动态网络预测蛋白复合物

目前蛋白质相互作用网络通常使用静态结构。但是,静态蛋白质网络是高度平均和理想化的网络模型,包含了在不同的条件、不同时间和不同空间发生的各种相互作用,而真实情况是随着外界条件的改变而发生变化的,蛋白质只有在特定的条件下,在相同的时间和相同的细胞位置才能发生相互作用完成某种生物过程。这个静态蛋白质网络没有办法体现出的动态特性。如何构建和分析动态蛋白网络是当前预测复合物的热点问题。

5) 构建多层网络预测蛋白质复合物

不同的生物信息可导致蛋白质之间不同的关系。目前现有的方法通常是将生物信息整合到蛋白质相互作用网络,各类生物信息具有相同的作用,并未考虑到不同类型的相互作用的影响。因而,未能有效地考虑不同生物信息具有不同的特性。因此,构建具有不同的生物信息的多层网络的结构会遵循自然特征,有益于发现更多具有生物功能的蛋白质复合物。

6) 大数据应用

随着高通量技术的发展,产生大量的相互作用数据以及多源的生物其他数据,这使源数据和数据库存在异质性,导致数据缺失、数据矛盾等问题,因而需要依靠大数据思维 and 数据分析技术对蛋白质相互作用多源数据进行深入挖掘和融合,进而有效预测蛋白质复合物。

致 谢

辽宁省教育厅科技项目(L201605)。

参考文献

- [1] 汤希玮. 蛋白质复合物预测算法综述[J]. 长沙大学学报, 2017, 31(5): 19-23.
- [2] 路东方, 许俊富, 项超娟, 谢江. 生物大数据中的聚类方法分析[J]. 上海大学学报(自然科学版), 2016, 22(1), 45-57.
- [3] Vlasblom, J. and Wodak, S.J. (2009) Markov Clustering Versus Affinity Propagation for the Partitioning of Protein Interaction Graphs. *BMC Bioinformatics*, **10**, 1-14. <https://doi.org/10.1186/1471-2105-10-99>
- [4] Dunn, R., Dudbridge, F. and Sanderson, C.M. (2005) The Use of Edge-Betweenness Clustering to Investigate Biological Function in Protein Interaction Networks. *BMC Bioinformatics*, **6**, 1-14. <https://doi.org/10.1186/1471-2105-6-39>

- [5] Bader, G.D. and Hogue, C.W. (2003) An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinformatics*, **4**, 2. <https://doi.org/10.1186/1471-2105-4-2>
- [6] Adamcsek, B., Palla, G., Farkas, I.J., I, D. and Vicsek, T. (2006) Cfinder: Locating Cliques and Overlapping Modules in Biological Networks. *Bioinformatics*, **22**, 1021-1023. <https://doi.org/10.1093/bioinformatics/btl039>
- [7] Altaf-Ul-Amin, M., Shinbo, Y., Mihara, K., Kurokawa, K. and Kanaya, S. (2006) Development and Implementation of an Algorithm for Detection of Protein Complexes in Large Interaction Networks. *BMC Bioinformatics*, **7**, 1-13. <https://doi.org/10.1186/1471-2105-7-207>
- [8] Chua, H.N.A.N.K., Sung, W.K., Leong, H.W. and Wong, L. (2008) Using Indirect Protein-Protein Interactions for Protein Complex Prediction. *Journal of Bioinformatics and Computational Biology*, **6**, 435-466. <https://doi.org/10.1142/S0219720008003497>
- [9] Abdullah, A., Deris, S., Hashim, S.Z.M. and Jamil, H.M. (2009) Graph Partitioning Method for Functional Module Detections of Protein Interaction Network. *International Conference on Computer Technology and Development, ICCTD'09*, Kota Kinabalu, 13-15 November 2009, 230-234.
- [10] Ramadan, E., Osgood, C. and Pothén, A. (2010) Discovering Overlapping Modules and Bridge Proteins in Proteomic Networks. *ACM International Conference on Bioinformatics and Computational Biology*, Niagara Falls, New York, 2-4 August 2010, 366-369.
- [11] Rhrissorakkrai, K. (2011) Mine: Module Identification in Networks. *BMC Bioinformatics*, **12**, 192. <https://doi.org/10.1186/1471-2105-12-192>
- [12] Efimov, D., Zaki, N. and Berenguères, J. (2012) Detecting Protein Complexes from Noisy Protein Interaction Data. 1-7. <https://doi.org/10.1145/2350176.2350177>
- [13] Wang, J.X., Li, M. and Ren, J. (2013) Identifying Protein Complexes Based on Density and Modularity in Protein-Protein Interaction Network. *BMC Systems Biology*, **7**, S12.
- [14] Wang, Y.J. and Qian, X.N. (2017) Finding Low-Conductance Sets with Dense Interactions (FLCD) for Better Protein Complex Prediction. *ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, **11**, 537-538.
- [15] Newman, M.E.J., Barabási, A.-L. and Watts, D.J. (2006) *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, 2006.
- [16] Aldecoa, R. and Marín, I. (2010) Jerarca: Efficient Analysis of Complex Networks Using Hierarchical Clustering. *PLoS ONE*, **5**, e11585. <https://doi.org/10.1371/journal.pone.0011585>
- [17] Ahn, Y.Y., Bagrow, J.P. and Lehmann, S. (2010) Link Communities Reveal Multiscale Complexity in Networks. *Nature*, **466**, 761-764. <https://doi.org/10.1038/nature09182>
- [18] Min, W., Le, O.Y. and Li, X.L. (2016) Protein Complex Detection via Effective Integration of Base Clustering Solutions and Co-Complex Affinity Scores. *IEEE/ACM Transactions on Computational Biology & Bioinformatics*, **14**, 733-739.
- [19] Van Dongen, S. (2000) Graph Clustering by Flow Simulation. PhD Thesis, University of Utrecht, Utrecht.
- [20] Hwang, W., Cho, Y.R., Zhang, A. and Ramanathan, M. (2006) A Novel Functional Module Detection Algorithm for Protein-Protein Interaction Networks. *Algorithms for Molecular Biology*, **1**, 24. <https://doi.org/10.1186/1748-7188-1-24>
- [21] Hwang, W., Cho, Y.-R., Zhang, A. and Ramanathan, M. (2008) Cascade: A Novel Quasi All Paths-Based Network Analysis Algorithm for Clustering Biological Interactions. *BMC Bioinformatics*, **9**, 64. <https://doi.org/10.1186/1471-2105-9-64>
- [22] Yukeng Shih, S.P. (2012) Identifying Functional Modules in Interaction Networks through Overlapping Markov Clustering. *Bioinformatics*, **28**, i473. <https://doi.org/10.1093/bioinformatics/bts370>
- [23] Ochieng, P.J., Kusuma, W.A. and Haryanto, T. (2017) Detection of Protein Complex from Protein-Protein Interaction Network Using Markov Clustering. *Journal of Physics: Conference Series*, **835**, 1-13.
- [24] Kamp, C. and Christensen, K. (2005) Spectral Analysis of Protein-Protein Interactions in *Drosophila Melanogaster*. *Physical Review E*, **71**, Article ID: 041911.
- [25] Qin, G. and Gao, L. (2010) Spectral Clustering for Detecting Protein Complexes in Protein Protein Interaction (ppi) Networks. *Mathematical and Computer Modelling*, **52**, 2066-2074. <https://doi.org/10.1016/j.mcm.2010.06.015>
- [26] Inoue, K., Li, W. and Kurata, H. (2010) Diffusion Model Based Spectral Clustering for Protein-Protein Interaction Networks.
- [27] Sallim, J., Abdullah, R. and Khader, A.T. (2008) ACOPIN: An ACO Algorithm with TSP Approach for Clustering Proteins from Protein Interaction Network. *2nd UKSIM European Symposium on Computer Modeling and Simulation*,

- Liverpool, 8-10 September 2008, 203-208.
- [28] Lei, X., Ying, C., Wu, F.X. and Xu, J. (2015) Clustering PPI Data by Combining FA and SHC Method. *BMC Genomics*, **16**, S3.
- [29] Feng, J., Jiang, R. and Jiang, T. (2011) A Max-Flow-Based Approach to the Identification of Protein Complexes Using Protein Interaction and Microarray Data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **8**, 621-634.
- [30] Li, M., Wu, X., Wang, J. and Pan, Y. (2012) Towards the Identification of Protein Complexes and Functional Modules by Integrating PPI Network and Gene Expression Data. *BMC Bioinformatics*, **13**, 1-16. <https://doi.org/10.1186/1471-2105-13-109>
- [31] Keretsu, S. and Sarmah, R. (2016) Weighted Edge Based Clustering to Identify Protein Complexes in Protein-Protein Interaction Networks Incorporating Gene Expression Profile. *Computational Biology & Chemistry*, **65**, 69-79. <https://doi.org/10.1016/j.compbiolchem.2016.10.001>
- [32] Shen, X., Zhou, J., Jiang, X., Hu, X., He, T., Yang, J. and Xie, D. (2017) A Novel Identified Temporal Protein Complexes Strategy Inspired by Density-Distance and Brainstorming Process. *IEEE International Conference on Bioinformatics and Biomedicine*, Kansas City, 13-16 November 2017, 1269-1274.
- [33] Przulj, N., Wigle, D.A. and Jurisica, I. (2004) Functional Topology in a Network of Protein Interactions. *Bioinformatics*, **20**, 340-348. <https://doi.org/10.1093/bioinformatics/btg415>
- [34] Ji, J., Liu, Z., Zhang, A., Jiao, L. and Liu, C. (2012) Improved Ant Colony Optimization for Detecting Functional Modules in Protein-Protein Interaction Networks. In: *Information Computing and Applications*, Springer, Berlin, 404-413. https://doi.org/10.1007/978-3-642-34041-3_57
- [35] Ji, J., Liu, Z., Zhang, A., Jiao, L. and Liu, C. (2012) Ant Colony Optimization with Multi-Agent Evolution for Detecting Functional Modules in Protein-Protein Interaction Networks. In: *Information Computing and Applications*, Springer, Berlin, 445-453. https://doi.org/10.1007/978-3-642-34062-8_58
- [36] Yang, Z.H., Yu, F.Y., Lin, H.F. and Wang, J. (2014) Integrating PPI Datasets with the PPI Data from Biomedical Literature for Protein Complex Detection. *BMC Medical Genomics*, **7**, 1-13. <https://doi.org/10.1186/1755-8794-7-S2-S3>
- [37] Yu, F., Yang, Z., Hu, X., Sun, Y., Lin, H. and Wang, J. (2015) Protein Complex Detection in PPI Networks Based on Data Integration and Supervised Learning Method. *BMC Bioinformatics*, **16**, 1-9. <https://doi.org/10.1186/s12859-014-0430-y>
- [38] Hu, L. and Chan, K.C. (2015) A Density-Based Clustering Approach for Identifying Overlapping Protein Complexes with Functional Preferences. *BMC Bioinformatics*, **16**, 174. <https://doi.org/10.1186/s12859-015-0583-3>
- [39] Li, X. (2016) Identification of Protein Complexes from Multi-Relationship Protein Interaction Networks. *Human Genomics*, **10**, 17. <https://doi.org/10.1186/s40246-016-0069-z>
- [40] Zhou, H., Liu, J., Li, J. and Duan, W. (2017) A Density-Based Approach for Detecting Complexes in Weighted PPI Networks by Semantic Similarity. *PLoS ONE*, **12**, e0180570. <https://doi.org/10.1371/journal.pone.0180570>
- [41] Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciati, C.M., et al. (2002) Functional Organization of the Yeast Proteome by Systematic Analysis of Protein Complexes. *Nature*, **415**, 141-147. <https://doi.org/10.1038/415141a>
- [42] Leung, H.C.M., Xiang, Q., Yiu, S.M. and Chin, F.Y.L. (2009) Predicting Protein Complexes from PPI Data: A Core-Attachment Approach. *Journal of Computational Biology*, **16**, 133-144. <https://doi.org/10.1089/cmb.2008.01TT>
- [43] Wu, M., Li, X.L., Kwok, C.K. and Ng, S.K. (2009) A Core-Attachment Based Method to Detect Protein Complexes in PPI Networks. *BMC Bioinformatics*, **10**, 1-16. <https://doi.org/10.1186/1471-2105-10-169>
- [44] Ma, X. and Gao, L. (2012) Predicting Protein Complexes in Protein Interaction Networks Using a Core-Attachment Algorithm Based on Graph Communicability. *Information Sciences*, **189**, 233-254. <https://doi.org/10.1016/j.ins.2011.11.033>
- [45] Srihari, N.K. and Leong, H. (2010) Mcl-Caw: A Refinement of Mcl for Detecting Yeast Complexes from Weighted PPI Networks by Incorporating Core-Attachment Structure. *BMC Bioinformatics*, **11**, 504-512. <https://doi.org/10.1186/1471-2105-11-504>
- [46] Keretsu, S. and Sarmah, R. (2017) Identification of Protein Complexes in Protein-Protein Interaction Networks by Core-Attachment Approach Incorporating Gene Expression Profile. *International Journal of Bioinformatics Research & Applications*, **13**, 313. <https://doi.org/10.1504/IJBRA.2017.087382>
- [47] Andreopoulos, B., Winter, C., Labudde, D. and Schroeder, M. (2009) Triangle Network Motifs Predict Complexes by Complementing High-Error Interactomes with Structural Information. *BMC Bioinformatics*, **10**, 407-413. <https://doi.org/10.1186/1471-2105-10-196>
- [48] Shi, L., Lei, X. and Zhang, A. (2011) Protein Complex Detection with Semi-Supervised Learning in Protein Interaction

Networks. *Proteome Science*, **9**, 1-9. <https://doi.org/10.1186/1477-5956-9-S1-S5>

- [49] Liu, Q., Song, J. and Li, J. (2016) Using Contrast Patterns between True Complexes and Random Subgraphs in PPI Networks to Predict Unknown Protein Complexes. *Scientific Reports*, **6**, Article No. 21223. <https://doi.org/10.1038/srep21223>
- [50] Yu, Y., Liu, J., Feng, N., Song, B. and Zheng, Z. (2017). Combining Sequence and Gene Ontology for Protein Module Detection in the Weighted Network. *Journal of Theoretical Biology*, **412**, 107-112. <https://doi.org/10.1016/j.jtbi.2016.10.010>

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2286, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: sea@hanspub.org