



蛋白质网络最短路径的查找

第七组：杜昊雨、张子谦、刘岚、刘胤辅

2022.5

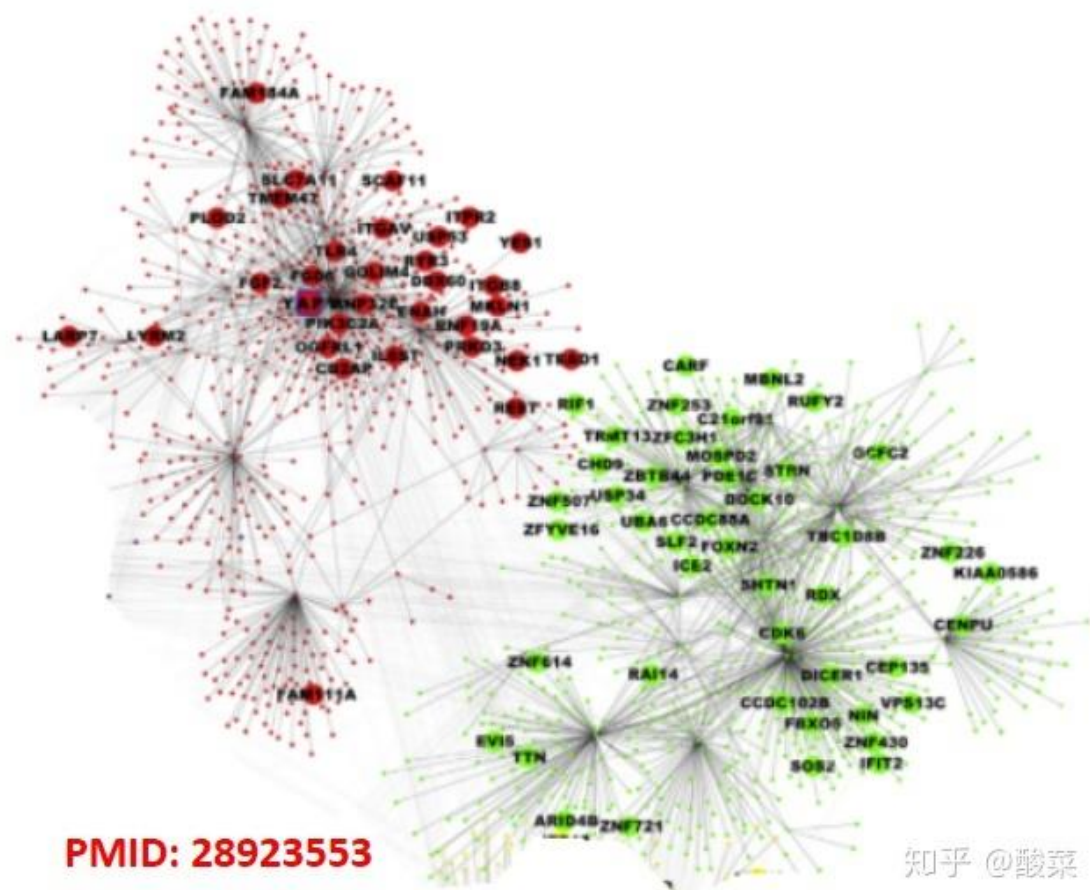


上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

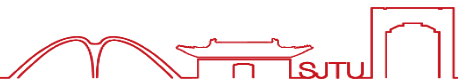
背景介绍

我们组的选题是：蛋白质网络的构建与最短路径的查询

蛋白质互作网络（Protein-Protein Interaction Networks, PPI）是由蛋白通过彼此之间的相互作用构成，来参与生物信号传递、基因表达调节、能量和物质代谢及细胞周期调控等生命过程的各个环节。系统分析大量蛋白在生物系统中的相互作用关系，对了解生物系统中蛋白质的工作原理，了解疾病等特殊生理状态下生物信号和能量物质代谢的反应机制，以及了解蛋白之间的功能联系都有重要意义。在蛋白质互作网络中，两个蛋白质之间的最短距离可能代表着他们之间的生物学联系；因此通过该网络中所得到的信息可以对所得到的蛋白进行一些生物信息方向的分析。



研究方案及分工

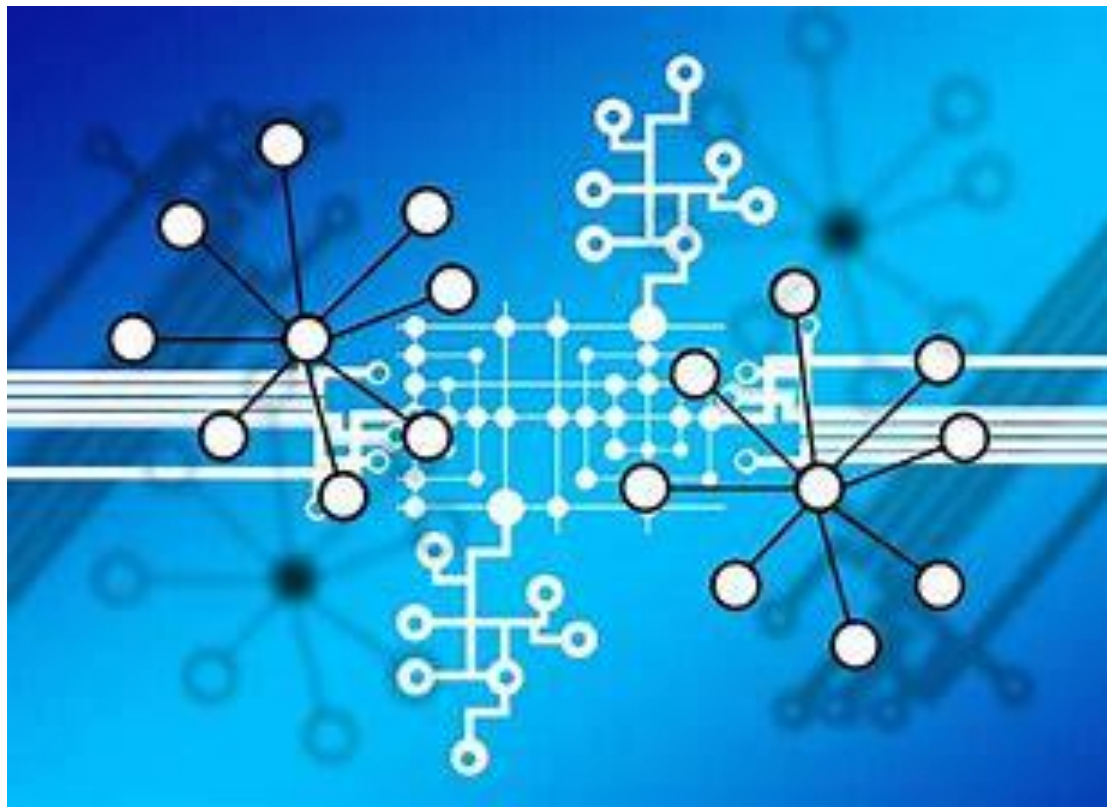


通过对相关资料的查找，我们已经知道对于最短路径的求法存在多种方案。首先，我们可能通过交互输入或是文件解析的方式引入蛋白质相关数据，之后，我们尝试构建一个蛋白质的网络。然后，我们尝试通过多种方式求得网络中每一个蛋白质到其他蛋白质的最短路径。

1. Floyd-Warshall Algorithm

2. Dijkstra algorithm

3. Bellman-Ford*algorithm



1. Floyd-Warshall Algorithm



Floyd算法是一个经典的动态规划算法，从任意节点*i*到任意节点*j*的最短路径不外乎2种可能，1是直接从*i*到*j*，2是从*i*经过若干个节点*k*到*j*。所以，我们假设Dis(*i,j*)为节点*u*到节点*v*的最短路径的距离，对于每一个节点*k*，我们检查Dis(*i,k*) + Dis(*k,j*) < Dis(*i,j*)是否成立，如果成立，证明从*i*到*k*再到*j*的路径比*i*直接到*j*的路径短，我们便设置Dis(*i,j*) = Dis(*i,k*) + Dis(*k,j*)，这样一来，当我们遍历完所有节点*k*，Dis(*i,j*)中记录的便是*i*到*j*的最短路径的距离。

这部分由刘胤辅负责。

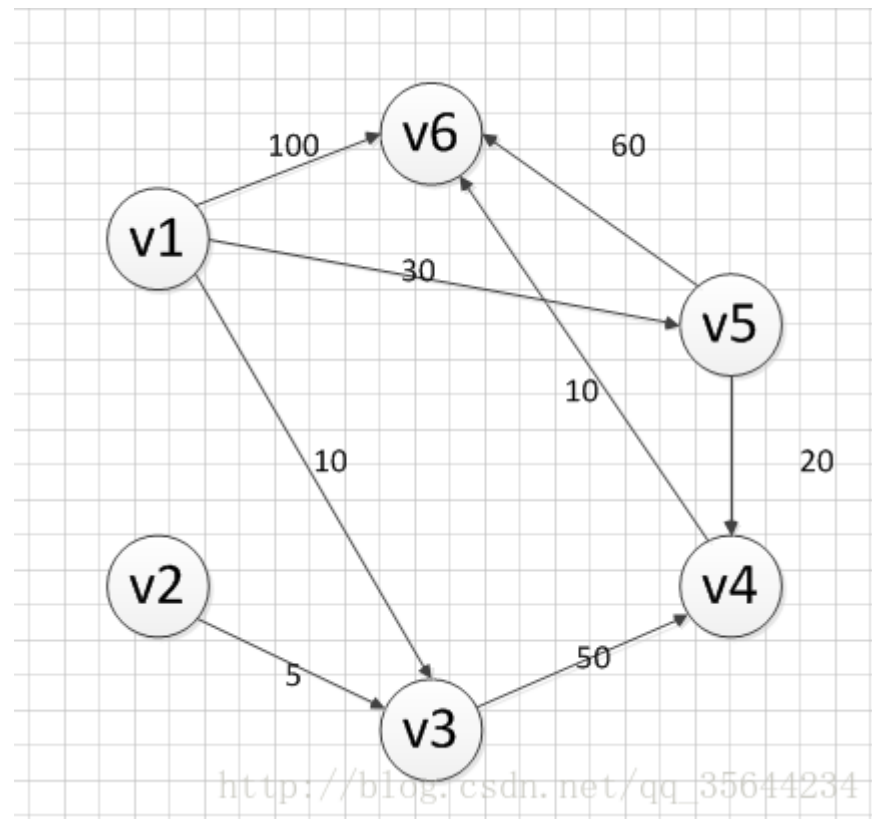
$$w_{ij}^{(0)} = \begin{cases} w_{ij} & \text{若 } e_{ij} \in E \text{ (有边)} \\ \infty & \text{若 } e_{ij} \notin E \text{ (无边)} \\ 0 & \text{若 } i = j \text{ (对角线元素)} \end{cases}$$

$$r_{ij}^{(0)} = \begin{cases} j & \text{若 } w_{ij}^{(0)} \neq \infty \\ 0, & \text{其它} \end{cases}$$

2. Dijkstra algorithm

迪杰斯特拉算法，是从一个顶点到其余各顶点的最短路径算法，解决的是有权图中最短路径问题。迪杰斯特拉算法主要特点是从起始点开始，采用贪心算法的策略，每次遍历到始点距离最近且未访问过的顶点的邻接节点，直到扩展到终点为止。

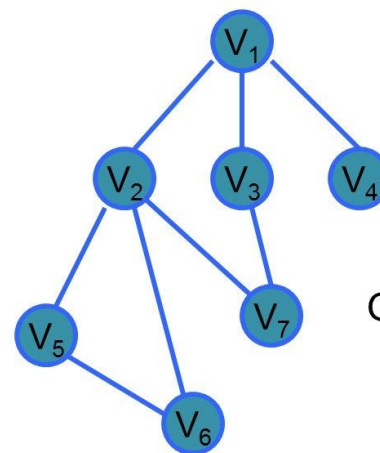
这个部分由张子谦负责。



3. Bellman-Ford algorithm

贝尔曼-福特算法，它的原理是对图进行次松弛操作，得到所有可能的最短路径。其优于迪科斯彻算法的方面是边的权值可以为负数、实现简单，缺点是时间复杂度过高，高达 $O(n^3)$ 。但算法可以进行若干种优化，提高了效率。贝尔曼-福特算法与迪科斯彻算法类似，都以松弛操作为基础，即估计的最短路径值渐渐地被更加准确的值替代，直至得到最优解。在两个算法中，计算时每个边之间的估计距离值都比真实值大，并且被新找到路径的最小长度替代。

这个部分由刘岚负责。



G.vexs=

V ₁	V ₂	V ₃	V ₄	V ₅	V ₆	V ₇
0	1	2	3	4	5	6

G.arcs=

	0	1	2	3	4	5	6
0	0	1	1	1	0	0	0
1	1	0	0	0	1	1	1
2	1	0	0	0	0	0	1
3	1	0	0	0	0	0	0
4	0	1	0	0	0	1	0
5	0	1	0	0	1	0	0
6	0	1	1	0	0	0	0

存放顶点的数组

表示边的矩阵

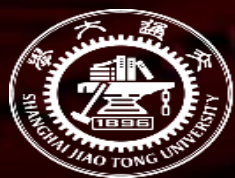
知乎 @承志

4. 数据输出、后期调试等



此部分由杜昊雨负责。

谢谢!



上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

