各位老师好，我是网络工程的张颖颖，我的毕设题目是基于异质网络分析的关键长非编码RNA的计算与识别

我的指导老师是朱剑林老师，和杜朴风老师

下面将从四个部分来介绍

第一部分研究背景

随着人类基因组计划的完成，人们发现大约有1.5%的负责编码蛋白质基因，剩下的将会转录成非编码RNA

其中的长非编码RNA是一类长度大于200个核苷酸且几乎没有蛋白质编码能力的非编码RNA。

前人对长非编码RNA的研究主要是识别预测、序列特征及结构预测、功能注释和与其他生物分子或疾病的关联预测等4个方面

长非编码RNA中也有着关键长非编码RNA

是指功能非常重要的一部分长非编码RNA

第二部分研究方法与思路

分为4步

首先利用lncRNA-protein数据集构建相应的异质信息网络

然后在利用HeteSim算法计算lncRNA之间的关联度

先选择lncRNA-protein-lncRNA作为元路径

再将关联关系转化邻接矩阵

再将矩阵分别按照行向量和列向量进行标准化得到2个转移概率矩阵

再进行标准化处理

就这样得到了lncRNA-lncRNA关联得分矩阵

第三步是利用上一步的关联得分矩阵构建lncRNA-lncRNA网络

利用复杂网络的拓扑结构，使用网络节点中心性来识别关键lncRNA

第四步进行实验结果的评估

因为无关键lncRNA数据库，所以采用GIC分数作为参照，来验证实验的有效性

然后将各个中心性计算出来的值按照从大到小的顺序进行排序

分别取前100，200，300，400作为识别出来的关键lncRNA

GIC分数也取5个阈值来区别关键lncRNA

最后采用灵敏度、特异性等7个指标来评估实验结果

第三部分主要难点与结论

先介绍4个难点

首先是无直接的lncRNA-lncRNA关联关系

使用heteSim算法在lncRNA-protein异质网络中计算出了lncRNA与其他lncRNA的间接关联数据

然后是如何识别关键lncRNA

使用lncRNA之间的关联数据建立对应的网络

在网络中认为节点越重要则越有可能是关键lncRNA

所以采用网络节点中心性的方式来识别

其次是各个网络节点中心性计算出来的值如何设定阈值

就是将各个值进行排序，取前100，200，300，400作为识别出来的关键lncRNA

最后是无关键lncRNA数据库，然后采用GIC分数来代替

GIC分数越高，则重要性越高，表明越有可能是关键lncRNA

再来介绍结论

GIC分数分别取了5个阈值，中心性阈值取Top100，计算了灵敏性、特异性等指标，如图所示

在图中，加粗的字体表示，在该组中的最高值，也就是效果最好的

我们可以发现，大部分情况下，BC中心性的识别效果更好

这5张图展示了不同阈值不同中心性下的ROC曲线，可以看出实验效果并不是很好，

因为AUC的值都在0.5左右，只有一个值较为突出为0.6674.

是在阈值为0.75，中心性为BC时，效果最佳

第四部分总结

首先介绍不足之处

有四点

前面2点是数据集和元路径都过于单一

第3点，用网络中心性识别关键lncRNA效果不是很好

第四点，本实验无真实准确的关键lncRNA数据库来判断识别结果的准确性，只能采用GIC分数作为代替，这也是导致实验效果不好的原因之一

然后介绍一下创新点

第1，采用hetesim算法在异质信息网络中计算出了lncRNA之间的间接关联关系

第2，基于复杂网络的拓扑结构，采用网络节点中心性的方法识别lncRNA-lncRNA网络中的关键lncRNA

第3，采用GIC分数作为是否为关键lncRNA的参照，弥补无关键lncRNA数据库的缺憾

好，我的讲解到此结束，谢谢大家

问题：

1. lncRNA与蛋白质有啥关系？

基因可以编码DNA片段，这些片段转录成RNA，就可以翻译成蛋白质，然后也有不编码的DNA，也可以转录，但是得到了RNA不能翻译成蛋白质，就是lncRNA。

lncRNA不能编码蛋白质，但是可以通过比如干扰其他基因的表达过程来调控对应的蛋白质。

2.介绍各个网络节点中心性？

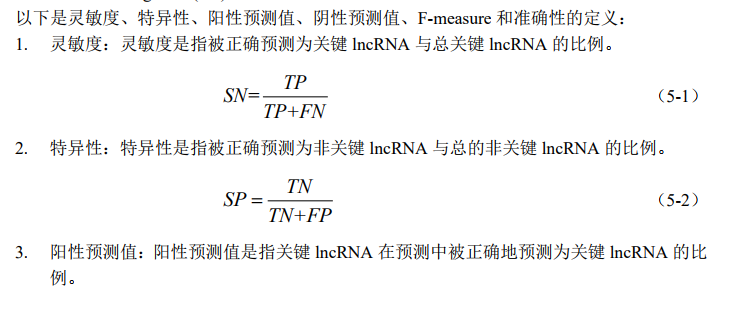
**紧密中心性：**这个点到其他的所有点的距离的平均值的倒数

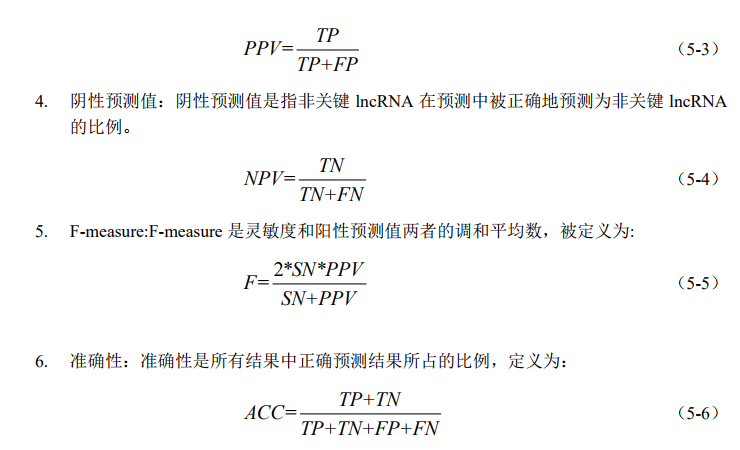
**介数中心性：**定义为通过该节点的最短路径的平均分数，就是节点s到节点t并且经过节点u的所有最短路径的数目。

**子图中心性：**度量了节点u所参与的整个网络的子图的数量。

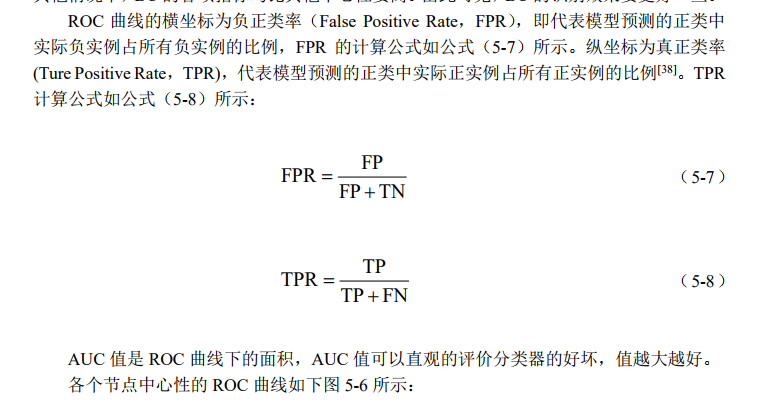
**特征向量中心性：**特征向量的第u个分量

**3、**灵敏性、特性性、。。。准确性？





1. ROC曲线？



1. HeteSim算法？

HeteSim算法是一种基于路径约束的相关性计算方法，认为2个对象之间的关联度与连接2个对象之间的路径有关。