# 使用python语言（用networkx包）

**也可以用Pajek软件，matlab.**

# 数据说明：

Name:代表用户名

Id: 代表理财产品编号

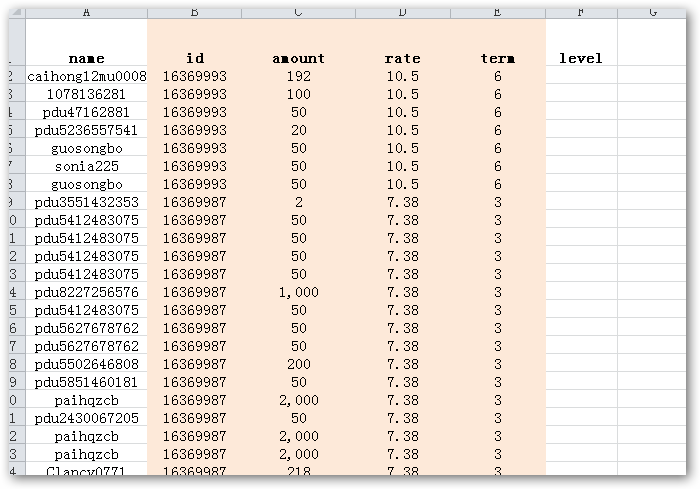
每个理财产品包括4个属性：

1.amount代表金额

2. Rate代表利息

3. Term代表周期

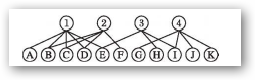
4.level代表这个理财产品的风险等级



# 二分网络拓扑结构

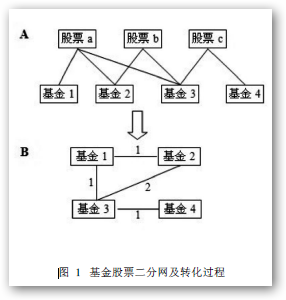
1. **构建“用户----产品”二分网**

有向 （就是带箭头的，箭头是从用户指向产品）的二分网络图，两类节点分别用不同颜色表示。不同排成两行，位置错综复杂可以。同类顶点之间没有联系, 即用户与用户之间、产品与产品之间没有关联, 如图所示



1. **调用投影函数，对用户-产品二分网进行加权投影处理，分别形成单顶点的用户网和产品网。**

将用户与产品之间的关系投影为产品之间相互关系的单顶点网络,如果两个用户购买了相同的产品, 那么这两个产品就存在相连的边, 权重定义为两个同类节点共同连接另一个节点的个数, 通过这种关系构建出“用户----产品”无向加权网络，用个基金-股票的例子，如下图：

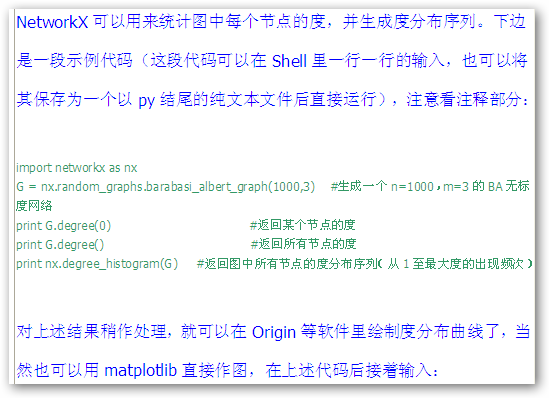


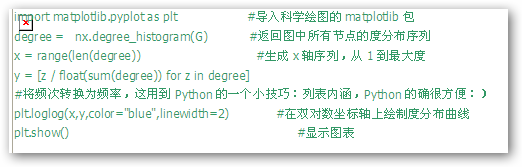
1. **对上面构建的两个单顶点网络进行统计计算**
2. **单顶点用户网络的gfl。**
3. **单顶点产品网络的平均度、最短路径长度和聚类系数统计值。**

这些统计值NETWORKX包中有相应的函数可以调用，可以看下帮助文件。

NetworkX提供了很多统计指标计算方法，可以查NetworkX的在线帮助文档：

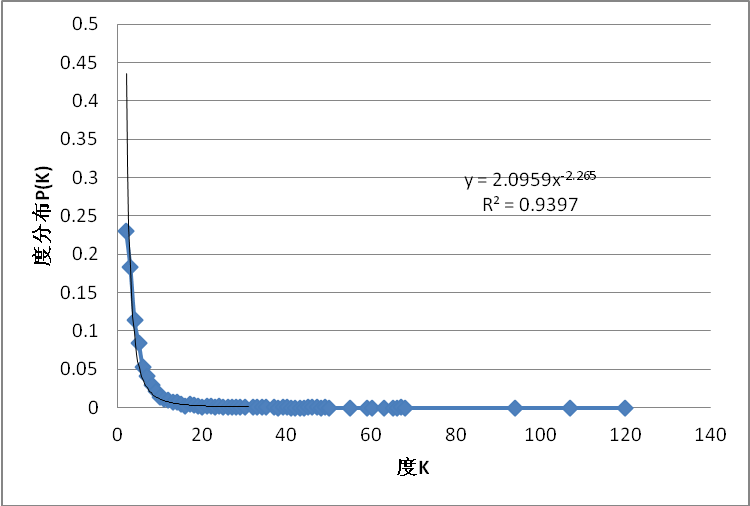
http://networkx.lanl.gov/reference/index.html



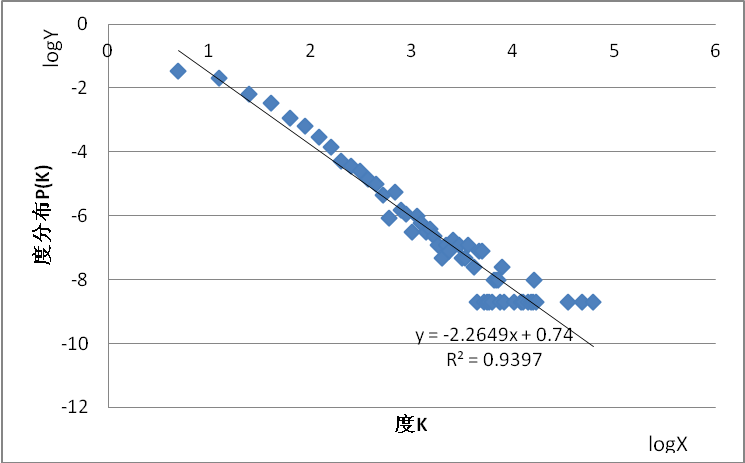


（一）度和度分布

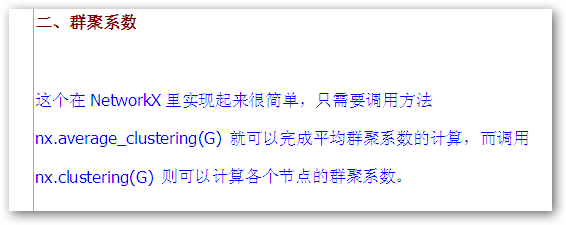
1. **对网络中所有节点的度求平均，得到网络的平均度，即把所有节点的度相加再除以节点数， 得到一个数值。**
2. **画出网络度分布图：横坐标为度K，纵坐标为节点的度为K的概率。为网络中度为的节点在整个网络中所占的比率，也就是，在网络中随机抽取到度为的节点的概率为，**



1. **进一步再把上图画在双对数坐标下：类似下图**



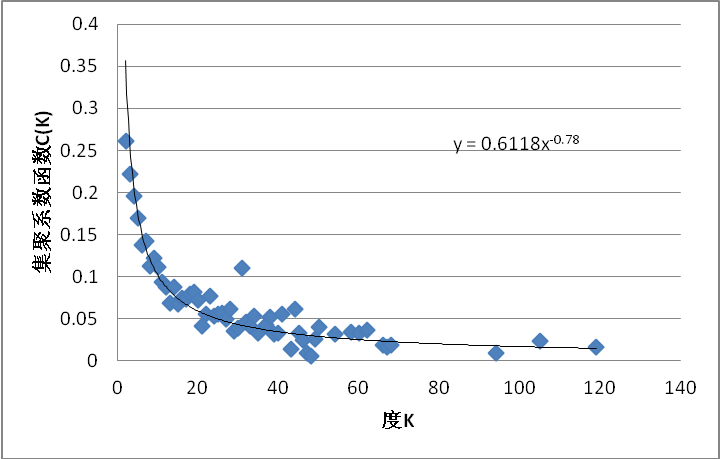
（二）集聚系数与聚度相关性



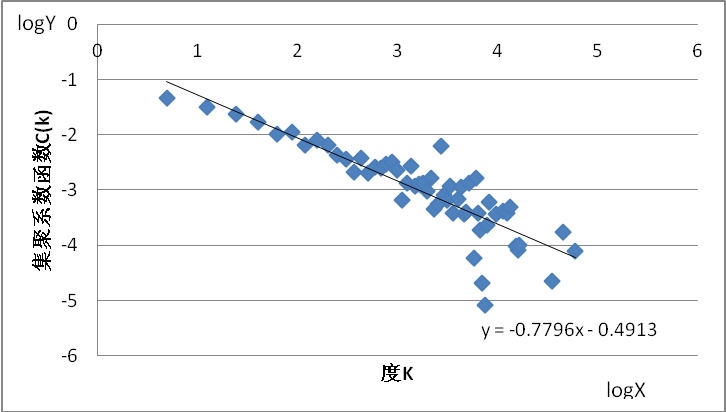
**1.计算得到网络的平均集聚系数C, 计算的是网络中与同一个节点连接的两节点之间也相互连接的平均概率。得到一个数值。**

**2. 计算具有相同规模的随机网络的平均聚集系数。。得到一个数值。**

**3. 在求得各节点集聚系数的基础上，可以得到度为的节点的集聚系数的平均值，绘制下方统计图**



4. 进一步再把上图画在双对数坐标下：类似下图：



**（二）网络平均最短路径**

1. **计算得到网络的平均最短路径L, 两节点之间的最短路径为：节点A到节点B所要经历的边的最小数目。网络的平均距离为所有节点对之间距离的平均值。得到一个数值。**

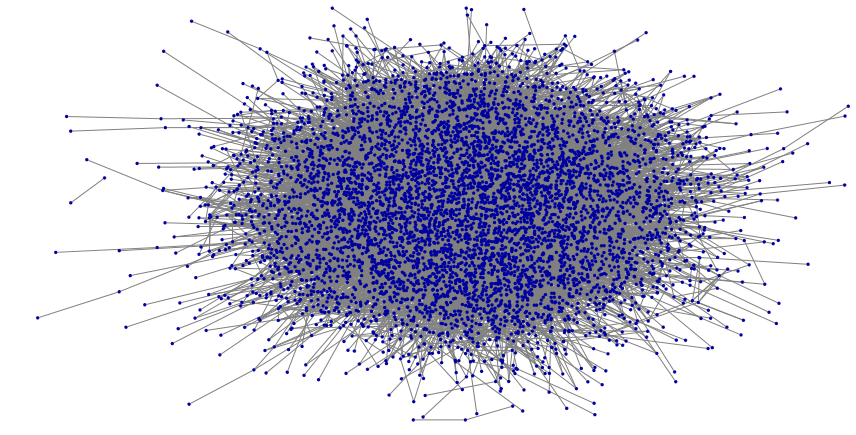
**2. 计算具有相同规模的随机网络的平均最短路径。得到一个数值。**

# 用户兴趣研究

**在构建“用户----产品”二分网基础上，绘制4个图：**

1. 横坐标为产品的amount（代表金额），纵坐标为产品的度。
2. 横坐标为产品的Rate (代表利息)，纵坐标为产品的度。
3. 横坐标为产品的Term (代表周期)，纵坐标为产品的度。
4. 横坐标为产品的level (代表产品的风险等级)，纵坐标为产品的度。

**构建“用户”有向网,类似下面这样的一个图，但是需要有出度入度箭头。**

****

**在上图基础上，将二分网中的用户分为两类，（1. 只有出箭头的用户 2. 既有出箭头又有入箭头的用户）**

分别对两类用户，绘制以上4个图：

1. 横坐标为产品的amount（代表金额），纵坐标为产品的度。
2. 横坐标为产品的Rate (代表利息)，纵坐标为产品的度。
3. 横坐标为产品的Term (代表周期)，纵坐标为产品的度。
4. 横坐标为产品的level (代表产品的风险等级)，纵坐标为产品的度。