山东大学 软件 学院

众智科学与网络化产业 课程实验报告

|  |  |  |  |
| --- | --- | --- | --- |
| 学号：202022300310 | 姓名： 张博宁 | | 班级： 软件20.7 |
| 实验题目：计算聚集系数和邻里重叠度 | | | |
| 实验学时：3 | | 实验日期： 2022/6/10 | |
| 实验目的：  课程实验可以验证、巩固和补充课堂讲授的理论知识，是《众智网络和网络化产业》课程中一个重要的、不可或缺的实践环节。其目的是通过实验，使学生能够验证课程中所传授理论，并将理论联系实际解决问题。本课程本身就是一门交叉课程，课程实验可以训练学生利用本专业知识和所学习理论，综合运用经济学、社会学、计算与信息科学以及应用数学的有关概念与方法，讨论社会学和经济学中的计算思维，讨论各类互联网时代众智现象背后的内在机理或规律，培养学生用理工科的方法与工具来分析社会科学和经济学等问题，开阔学生视野，更好地面对未来的挑战。 | | | |
| 硬件环境：  处理器 AMD Ryzen 7 5700U with Radeon Graphics 1.80 GHz  机带 RAM 16.0 GB  系统类型 64 位操作系统, 基于 x64 的处理器 | | | |
| 软件环境：  Windows 11 家庭中文版  Embarcadero Dev-C++ Version 6.3  TDM-GCC 9.2.0 64-bit Release  Gephi 0.9 | | | |
| 实验步骤与内容：  **1、实验要求：**  输入：任意的无向图  输出：  1）每个节点的聚集系数  2）每个节点对的邻里重叠度  **2、分析**  原题目为有向图，然而以下两种属性为无向图属性，且经询问后，确认题目应为无向图。  根据定义直接计算即可：  聚集系数：节点A的聚集系数 = A的任意两个朋友之间也是朋友的概率（即邻居间朋友对的个数除以总对数）  邻里重叠度：与A、B均为邻居的节点数/ 与节点A、B中至少一个为邻居的节点数  **3、代码编写**  使用邻接矩阵读入一个无向图后，对每个元素进行如下计算：  聚集系数：   1. for(int i=1;i<=n;i++){ 2. int cnt=0,s=0; 3. for(int j=1;j<=n;j++)if(i!=j&&g[i][j]){ 4. cnt++; 5. for(int k=j+1;k<=n;k++)if(g[j][k]&&g[i][k]&&k!=i)s++; 6. } 7. if(!cnt||cnt==1)continue; 8. clu[i]=((double)s/(double)(cnt\*(cnt-1)/2)); 9. }   使用比较原始的枚举暴力方法，枚举每个点，直接根据定义计算聚集系数，共n个节点，计算每个节点的聚集系数复杂度为O(n^2)，总时间复杂度为O(n^3)  邻里重叠度：   1. for(int i=1;i<n;i++) 2. for(int j=i+1;j<=n;j++){ 3. int cnt=0,s=0; 4. for(int k=1;k<=n;k++)if(k!=i&&k!=j){ 5. if(g[j][k]||g[i][k])cnt++; 6. if(g[j][k]&&g[i][k])s++; 7. } 8. nei[i][j]=nei[j][i]=(double)s/(double)cnt; 9. }   同样，直接枚举每一条边，根据定义直接计算邻里重叠度，邻接矩阵中枚举边的复杂度为O(n^2)，计算邻里重叠度的复杂度为O(n)，总复杂度依然是O(n^3)  此程序的输入输出部分不再赘述，输出效果见下文。  **4、调试与结果观察**  为了方便多组样例/大样例的观察，我编写了邻接矩阵生成器，代码如下：   1. for(int i=0;i<V;i++) 2. for(int j=0;j<i;j++) 3. if(rand()%100<CRD\*100)G[i][j]=G[j][i]=1;   其中CRD为与稠密度相关的0~1实数类型参数，生成的思路即每个节点对之间，有CRD的概率存在一条边。  下图为CRD=0.4时，生成的无向图：    根据此图的邻接矩阵计算的结果如下：    可以发现，与人工计算的结果一致，基本可以说明程序正确。  另外根据数据，我们可以验证聚集系数和邻里重叠度的意义：  聚集系数即节点的凝聚力表现。可见图中邻居较少的节点拥有更高的邻里重叠度（或者说更极端的邻里重叠度），而邻居多的节点如7、5、6，拥有与“稠密度”相似的邻里重叠度，反映了图的整情况。  邻里重叠度表示了一个边的“捷径”程度，即社会网络中判断某点是否能够通过某条边更便利的到达另一点时对该边的属性的刻画。图中7-5的边邻里重叠度较大，所以作为捷径的可能性较小，也可以发现该边的两个端点拥有较多的邻居。  而图中9-10边的邻里重叠度较小，更可能作为捷径。 | | | |
| 结论分析与体会：  经过本次实验，我更加熟练掌握了聚集系数和邻里重叠度的定义，更熟练了模拟算法的使用与对浮点数的操作。  此外，经过数据的分析过程，我了解了开源软件Gephi的使用。对于结果的分析让我更了解了“捷径”的概念与节点的“凝聚力”，他们可以用本实验计算出的两个属性值进行量化。 | | | |

目录下有generator.cpp与nei&clu.cpp

编译后，运行generator.exe可在其目录中生成data.txt，内容为邻接矩阵

此时运行nei&clu.exe即可显示结果