

### 3.1 Ucinet 软件介绍

#### 3.1.1 发展简史

Ucinet 软件是由加州大学欧文分校的马克·林博特 (Mark E. J. Borgatti) 和史蒂夫·安德烈 (Stephen A. Brandes) 开发的。这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。林博特和安德烈在 1995 年开发了这款软件，并将其命名为 Ucinet。这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。林博特和安德烈在 1995 年开发了这款软件，并将其命名为 Ucinet。这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。



# 第3章

## 当代最流行的 社会网络分析软件 Ucinet

这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。林博特和安德烈在 1995 年开发了这款软件，并将其命名为 Ucinet。这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。林博特和安德烈在 1995 年开发了这款软件，并将其命名为 Ucinet。这款软件最初是为分析社交网络而设计的，后来被广泛应用于各种类型的网络分析。

## 3.1 Ucinet 软件介绍

### 3.1.1 发展简史

Ucinet 软件是由加州大学欧文 (Irvine) 分校的一群网络分析者编写的。现在对该软件进行扩展的团队由斯蒂芬·博加提 (Stephen Borgatti)、马丁·埃弗里特 (Martin Everett) 和林顿·弗里曼 (Linton Freeman) 组成。该软件最初是一组用 Basic 语言编写的模块, 逐渐发展成为综合性的 DOS 程序, 现在已经可作为一种 Windows 程序来使用了。这是一个具有通用目标、易于使用的程序, 它涵盖了一些基本的图论概念、位置分析法和多维量表分析法等。

#### 【知识扩展】加州大学欧文分校

加州大学欧文分校 (University of California-Irvine, UCI) 又称伊荣分校, 是一所四年制公立大学, 成立于 1965 年。尽管它是加州大学各成员学院中历史最短的一所大学, 但排名却紧追伯克莱、洛杉矶和戴维斯等大学之后, 列全美最好的公立学校前十名之一。加州大学欧文分校除了拥有崭新优美的建筑外, 教职人员皆具有深厚的专业背景, 其中有两人为诺贝尔奖获得者。UCI 最好的学科是生物科学、社会生态学、英文、经济和政治科学。其中, 生物科学是全美国大学最好和学生选修最多的学科。



#### 【人物小传】斯蒂芬·博加提

斯蒂芬·博加提 (Stephen Borgatti, 生于 1956 年) 在 1989 年从加州大学欧文分校获得数学社会学博士学位, 曾为美国波士顿学院 (Boston College) 卡罗尔管理学院 (Carroll School of Management) 教授、肯塔基大学 (University of Kentucky) 加顿商业与经济学院 (Gatton College of Business and Economics) 管理学教授, 其研究领域包括: 社会网络、知识管理、社会认知与结构、文化领域分析、正式角色理论等。

#### 【人物小传】马丁·埃弗里特

马丁·埃弗里特 (Martin G. Everett) 从牛津大学获得数学和社会学博士学位, 现为英国曼彻斯特大学 (University of Manchester) 社会科学学院 (School of Social Sciences) 教授。他参与创建了社会网络分析米切尔中心 (现改名为曼彻斯特社会网络组), 其主要研究兴趣在数学和社会学分析领域, 参加过隐蔽社会网络、精神病人网络、技术分析网络、自我网络、网络和扩散等多个项目的研究。他还是目前最常用的社会网络数据分析软件 Ucinet 的共同著作权人, 曾担任社会网络分析国际网络 (INSNA) 的总裁, 并在 2004 当选为社会科学院 (Academy of Social Sciences) 院士。



### 【人物小传】林顿·弗里曼

林顿·弗里曼 (Linton C. Freeman) 在 1956 年获得美国西北大学社会学博士学位, 现为加州大学欧文分校 (UCI) 社会学系和数理行为科学研究所的研究教授, 他一直致力于社会网络分析、社会结构和认知领域的研究, 并在 1978 年创办了“社会网络 (Social Networks)”杂志, 代表性著作有《社会网络分析发展史》, 是社会网络分析发展的推动者, 并为此做出了巨大贡献。社会网络分析研究领域的最高奖弗里曼奖就是用其名字来命名的。

新版本的 Ucinet 6.0 是基于 Ucinet 5.0 版本修改而成的, 所以熟悉使用 Ucinet 5.0 的用户可以轻松熟悉 Ucinet 6.0 的环境。新版本不仅拓展了软件的功能, 而且重新组织了程序的归类, 使其更加合理。软件最主要的设计思路是考虑应该提供哪些功能。自从第一个版本发布, Ucinet 一直吸收不同的网络算法。这些算法无论是在功能上 (检测具有凝聚力的群体、测量中心度等) 还是来源上 (不同人以不同的数学思想、方法论的观点下创造出来的) 都很多样。

Ucinet 软件提供了很多指导文档, 包括 Ucinet 快速上手指南、社会网络分析方法介绍手册、Ucinet 参考手册、Ucinet 用户指南以及在线帮助文档。

- Ucinet 快速上手指南提供了快速使用 Ucinet 的使用说明, 对软件的工作界面及数据处理流程做了简单的介绍, 可以帮助读者快速了解软件。
- 社会网络分析方法介绍是在线手册, 介绍了许多基本的、常用的社会网络分析方法。这本手册主要是由 Ucinet 软件的创始人 Freeman、Borgatti 和 Everett 撰写, 同时也得到加州大学洛杉矶分校的 Phillip Bonacich 教授指导的研讨班的帮助。
- Ucinet 参考手册介绍了 Ucinet 软件中所有菜单功能, 包括菜单功能的目的是、功能描述、参数设置、结果文件、功能使用主要事项等。
- Ucinet 用户指南具体介绍了 Ucinet 新版本特点、程序功能、数据形式等, 是为读者系统学习软件的使用而撰写的手册。

### 3.1.2 主要特点

Ucinet 是菜单驱动程序, 菜单是嵌套的, 选择菜单中的某个功能时可以显示该功能下的所有子菜单, 用户可根据需要选择。新版本的 Ucinet 6.0 与以往追求易用性相比, 更注重程序运行的速度。在运行 Ucinet 程序时, 需要在降低最大处理的网络结点大小、占用大量内存的算法, 还是使用能处理更多数据集合、节省内存空间的缓慢算法之间做出选择。Ucinet 6.0 版本中, 一般选择速度: ①当计算大量数据集时, 速度是必须的, 如果一个算法运行上千结点需要耗费几天的时间, 怎么能算一个好算法呢? ②计算机硬件及操作系统的快速升级扩展了程序运行的内存空间, 所以想办法设法做出节约内存的程序算法也就成为了浪费时间的事情。

Ucinet 采用菜单系统的后果之一, 是需要将程序功能组织成类别及子类, 并且要富于逻辑、可理解。事实证明, 这不是不可能的, 但 Ucinet 提供的所有功能分类有几套竞争方案, 但没有一个方案是完美的。每个方案在对一部分功能归类非常完美, 而对于剩



余部分功能的归类就显得很勉强。Ucinet 已经使用的分类方案同样存在这样的问题，但最基本的思想是在网络环境下，该方案是基于网络理论基础之上的。

Ucinet 网络分析集成软件包括一维与二维数据分析的 NetDraw，还有正在发展应用的三维展示分析软件 Mage 等，同时集成了 Pajek 用于大型网络分析的 Free 应用程序。利用 Ucinet 软件可以读取文本文件、KrackPlot、Pajek、Negopy、VNA 等格式的文件。它能处理 32767 个网络结点。当然，从实际操作来看，当结点数在 5000~10000 之间时，一些程序的运行就会很慢。

Ucinet 提供大量的数据管理和转换工具，如选择子集、合并数据集、序化、转化或记录数据。Ucinet 不包含可视化的过程，但是它与软件 Mage、NetDraw 和 Pajek 集成在一起，从而能够实现可视化。NetDraw 是简单的绘制网络图的工具，它可以读取 Ucinet 系统文件、Ucinet 文本文件、Pajek 文本文件等。它可以同时处理多种关系，并可以根据结点的特性设置结点的颜色、形状和大小，并可做数据分析，如中心性分析、子图分析、角色分析等，也具有很强的矩阵运算能力，是一个非常灵活的可视化软件。

Ucinet 下载后可免费使用两个月，试用版本的数据是过期的，功能上是和正式版本一样的。试用期满后可以在线购买，并针对用户类型规定了不同的购买价格。

### 3.1.3 运行环境

Ucinet 6.0 的运行环境需要 Windows 95 (从 1997 年或更近)、Windows 98、Windows NT 或其他兼容的操作系统，同时需要 2 MB 的硬盘存储空间和 16 MB 的缓冲内存。

Ucinet 必须安装程序才能够使用。如果把安装 CD 放入磁盘驱动器中，程序可以自动开始安装。如果有电子版本软件包，选择文件 setup.exe，安装向导将指导完成安装过程。若退出软件，可以选择 File | Exit 命令，或者单击退出按钮，或者使用快捷键【Ctrl+X】。如果软件正在执行分析功能，需要停止程序运行，可以按【Esc】键中断程序，这些操作一般适用于迭代程序 MDS 或 TABU Search 等。除此之外，还可以同时按【Ctrl+Alt+Delete】快捷键，调出操作系统的任务管理器窗口，单击“应用程序”按钮，选择 Ucinet 并结束任务。

Ucinet 软件是菜单驱动下的程序，使用起来很方便。启动软件后，可以发现主窗口有文件、数据、转换、工具、网络、可视化、选项和帮助菜单及几个快捷按钮（不同版本的菜单及按钮稍有不同）。打开每一个菜单，有更多的子菜单出现；快捷按钮有直接退出、电子表格编辑器、DL 数据语言编辑器、数据集调用、运行 Netdraw 软件等。

选择菜单功能的方式共有两种，最简单的方式是用鼠标单击需要选择的菜单，可以直接运行程序。如果所选择的菜单下面还有子菜单，子菜单会自动显示。另一种运行程序的方式是从键盘输入要选择菜单的快捷字母（一般是菜单的首字母），若选择的菜单没有子菜单时，软件就会开始运行所选择的功能。如若要选择 Network 菜单，可以在键盘上按 L 键就定位到了 Roles and Positions 功能，之后在展现出的子菜单功能中进一步选择时，可以按【S】键选中 Structural 功能，接着按【P】键可以选中 Profile 并运行该功能。

另外，有些功能可以不使用菜单，比如菜单中的 Data | Display 功能，可以直接使用预先设置好的快捷键激活相应程序直接运行。假设预先没有选定菜单，那么按快捷键



【Ctrl+D】将马上激活 Data | Display 功能。

所有快捷键都是【Ctrl】或者【Alt】键与其他字母的组合，如表 3-1 所示。

表 3-1 快捷键及其功能

快捷键	菜单功能
Ctrl+F	File   Change Default Folder
Alt+X	File   Exit
Ctrl+S	Data   Spreadsheet
Ctrl+D	Data   Display
Ctrl+B	Data   Describe
Ctrl+X	Data   Extract
Ctrl+M	Data   Random>Matrix

如果从菜单中选择一个路径，除少数命令例外，一般都会弹出参数表格。参数表格是一个或多个对话框的集合。比如，选择 Network | Subgroups | Cliques 命令后看到的表格形式如表 3-2 所示。

表 3-2 选择 Network | Subgroups | Cliques 命令后看到的表格形式

问题	默认选择
Input dataset:	Camp92.##h
Minimum size:	3
Analyze pattern of overlaps?:	Yes
Diagram Type	Tree Diagram
(Output) Clique indicator matrix:	Clique Sets
(Output) Co-membership matrix:	Clique Overlap
(Output) Partition indicator matrix:	Clique Part

上述问题表单中左侧的黑体字为问题，右侧显示的是默认的选择。通常用户会保留默认值，除了要输入的数据集，需要更改为自己要分析的数据集名称。更改数据集名称可以自己手动输入，也可以单击右侧按钮，出现选择文件列表对话框。表格中的问题确定后，即可单击 OK 按钮加载程序进行计算或分析。下次再选择菜单中的 Clique 路径时，弹出的参数表格中默认的 Dataset 集合是上次输入的数据集。

Ucinet 6.0 程序运行可以通过两种方式，并相应的产生两种形式的结果（见图 3-1）。

一种输入形式是 Ucinet 6.0 数据集。大多数程序接受一个单独的数据集作为输入，但可能是包含多个矩阵的数据集，这些矩阵通常代表着行动者不同的社会关系。其他程序（如 QAP 回归分析程序）接受多个数据集作为输入，每个数据集可能包含多个矩阵。

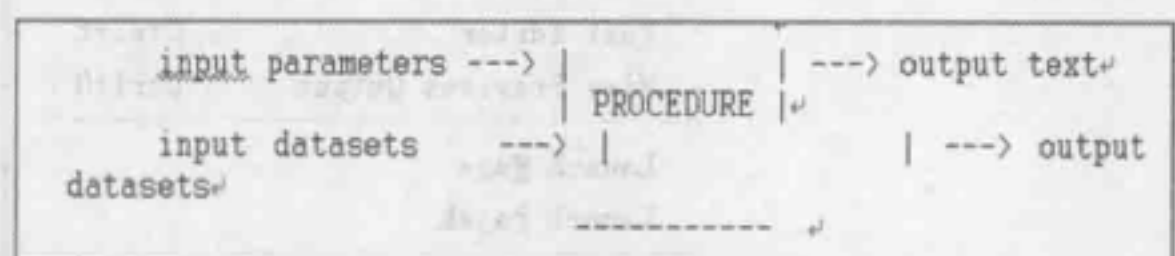


图 3-1 Ucinet 6.0 程序运行方式

另一种输入形式是输入参数描述数据类型或选择运行的程序。例如，如上面看到的，Clique 程序的参数输入是：输入数据集的名称，派系的最小规模，结果文件的名称。多数参数都有默认值，用户可以选择默认值也可以自行修改。

菜单中的某个路径一旦被选择后，表单中的参数就会被指定并立即出现。第一次调用程序的会话窗口中，参数都会有默认值，如果需要改变这些默认值，这些默认值会在接下来的运行过程中继续起作用，直到再修改它或者关闭 Ucinet 时才会失效。唯一的例外是，一些默认的设置依赖于其他默认设置的值，在这种情况下更改某些参数值将会改变其他程序的默认值，即使其他程序的默认值被提前设置也会被修改。

其中默认参数值一直不会改变的是程序运行后的输出结果文件名称，结果文件都是 Ucinet6.0 数据集，并且可以被 Ucinet 软件中其他的程序查看或者用来进行深层分析使用。例如，Clique 程序的输出文件是行动者矩阵，矩阵中的值给出的是派系的数目。这个矩阵可以输入 MDS 程序获得多维空间迭代分析，该数据集的默认输出名称是 Clique Overlap。

除了输出数据集，大多数 Ucinet 的过程也产生一个文本报告作为输出。这份报告总是保存到一个文本文件称为一个日志文件（通常存储在\Windows\系统目录），也会自动显示在屏幕上。

输出的文本文件存储在系统目录上，默认名称为 Log File Number X，X 是一个整数，用户可以利用工具栏的按钮修改、保存或者打印文本文件。

### 3.2 Ucinet 软件功能

#### 3.2.1 主要功能

Ucinet 主菜单中有许多功能菜单项，有些功能的含义相对较复杂，以下结合笔者的理解对 Ucinet 主菜单功能进行简单介绍，供读者参考。

##### 1. File 菜单

在 Ucinet 主界面选择 File 菜单，会出现子菜单，如图 3-2 所示，选择相应的命令，可以进行相应的数据处理和操作。

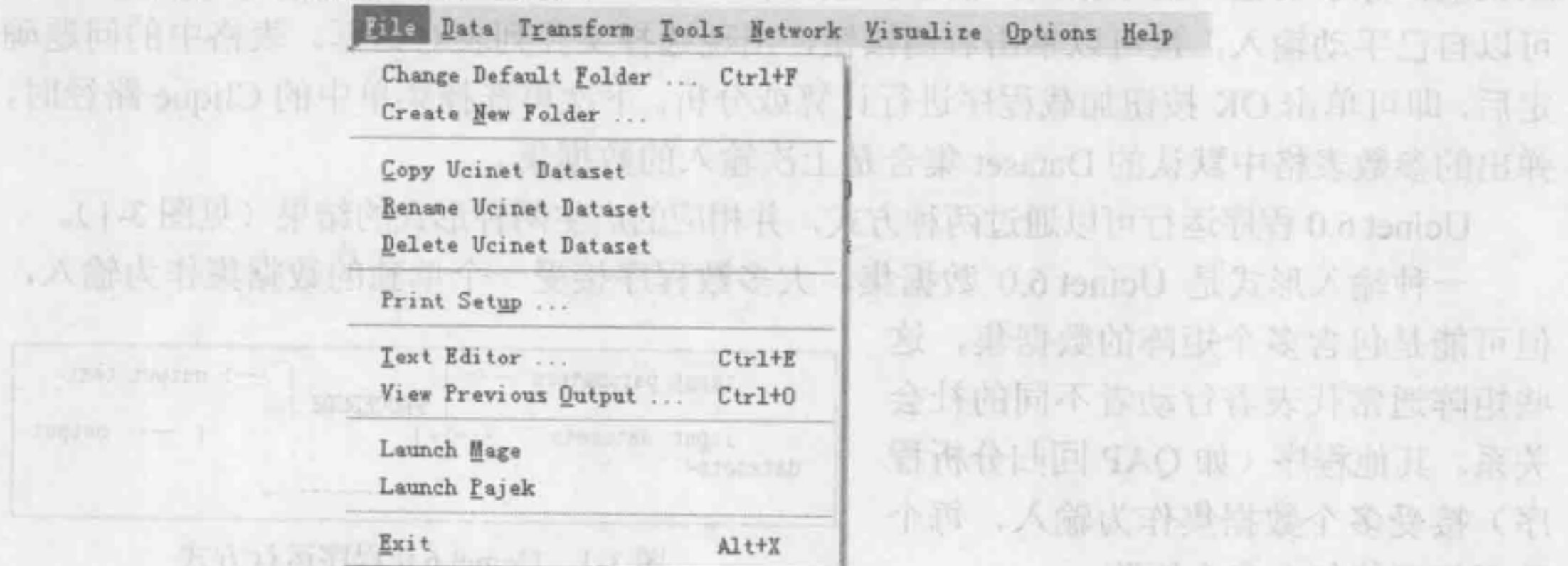


图 3-2 File 菜单

- Change Default Folder: 改变默认的文件夹。进行数据分析时，用户最好将自己



经常使用的文件夹设置成默认文件夹，以便于数据处理。Ucinet 都是在默认文件夹中搜索文件的，输入文件名后，Ucinet 都会在当前路径文件夹下进行搜索，除非输入的是文件的完整路径。如果要改变默认的文件夹，需要用到 Change Default Folder 命令，然后就可以选择经常使用的文件夹作为默认的文件夹。

- Create New Folder: 创建新文件夹。可以把经常使用的数据放在所创建的新文件夹中，只要打开 Ucinet，默认的文件夹就是新创建的文件夹。
- Copy Ucinet Dataset: 复制一个已经存在的 Ucinet 数据，再新建一个文件名。
- Rename Ucinet Dataset: 对已经存在的 Ucinet 数据集进行重新命名。
- Delete Ucinet Dataset: 删除已存在的 Ucinet 数据集。
- Print Setup: 设置打印机。
- Text Editor: 文本编辑器。用来输入、编辑、保存 Ucinet 数据文件。文本编辑器是创建 Ucinet 数据集的方式之一。
- View Previous Output: 查看在分析过程中产生的日志文件 (Log File)。
- Launch Mage: 开启 Mage 画图程序，进行可视化分析。
- Launch Pajek: 开始 Pajek 画图程序，进行可视化分析。
- Exit: 退出 Ucinet 程序。

## 2. Data 菜单

数据菜单 (Data) 中包含了处理 Ucinet 数据的命令，它对数据文件进行输入、编辑、输出以及显示分析的结果等，主要包括 6 类，如图 3-3 所示。

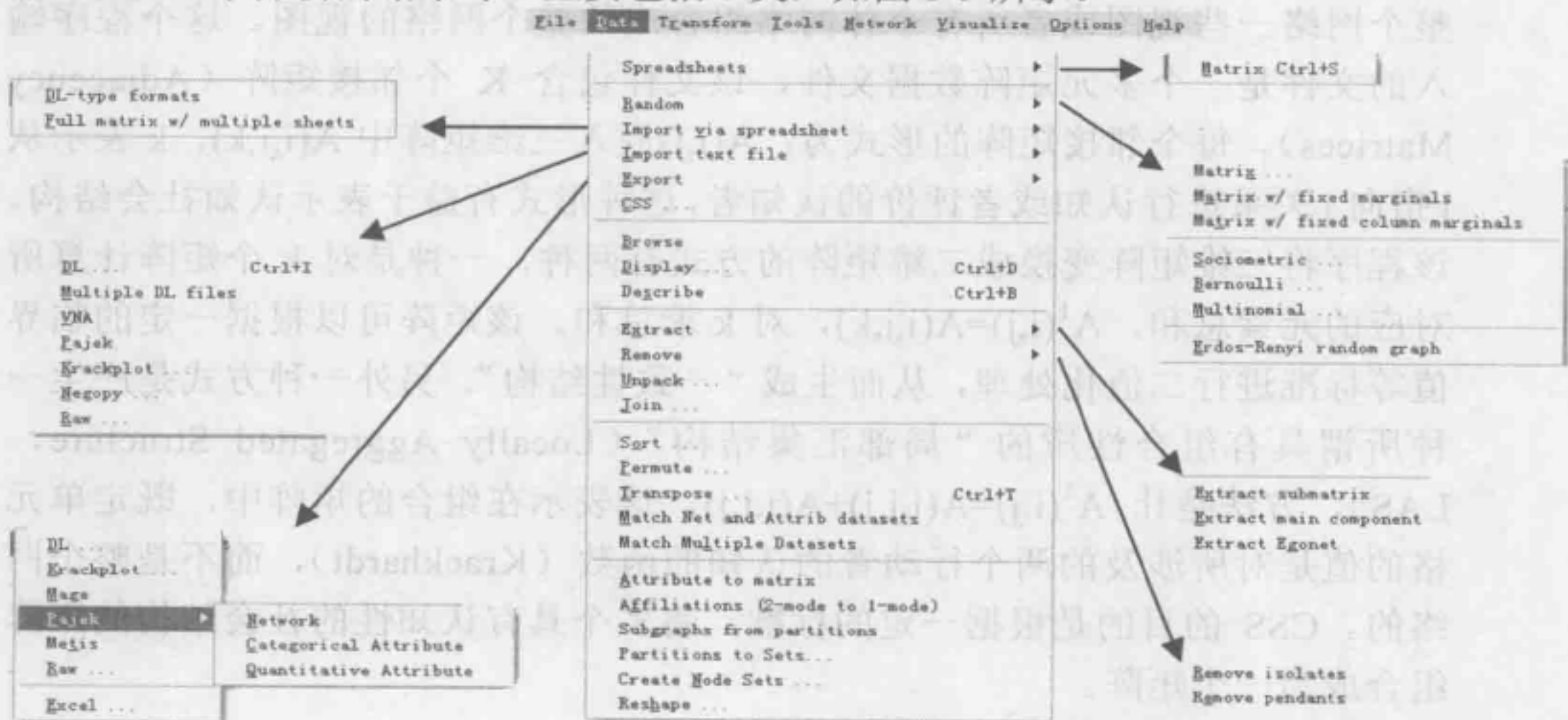


图 3-3 Data 菜单

### (1) Spreadsheets 操作

Spreadsheets (即 Ucinet 数据表编辑器) 的快捷键为【Ctrl+S】，可用它来输入、编辑、保存 Ucinet 数据集，可以打开已经存在的 Ucinet 数据集，对数据集的行列数据进行变换、插入、删除等操作，可以对新的数据表进行对称化、转置、二值化处理，或者输入随机数据等。

## (2) 数据的输入和输出等操作

具体包括以下 5 项内容。

- **Random:** 创建随机数据, 可创建的数据形式多种多样。如创建随机矩阵, 可以对矩阵需要的行列、矩阵的对角线等进行设置, 其格值都是满足一定概率分布的 Ucinet 数据; 还可以创建计量社会学随机图表以及伯努力分布矩阵和多项分布矩阵等。
- **Import via spreadsheet:** 利用 Spreadsheet 输入文件。该命令可以将 Excel 类型的文件转换为 Ucinet 数据, 也可以直接输入 DL 格式的数据文件进行编辑、保存等。
- **Import text file:** 输入文本文件、可输入 DL (Data Language) 格式文件及多元数据格式 (Multiple Data Files) 文件、VNA 软件使用的文件 (VNA 格式)、Pajek 软件使用的文件、Krack 制图软件格式的文件 (Krackplot)、Negopy 软件的文件格式、ASCII 型的初始文件 (Raw) 以及描述性的文本记录文件。
- **Export:** 输出文件。该项功能是将 Ucinet 产生的文件输出到其他格式类型的文件, 可以供其他软件分析使用。Ucinet 支持的输出文件格式与支持的输入文件格式基本一致, 包括数据语言文本文件格式 (DL)、Krackplot 文件、Mage 文件、Pajek 文件、Metis 文件、Raw 文件及 Excel 文件。
- **CSS:** 输入认知社会结构 (Cognitive Social Structure, CSS) 类型的文件, 其功能是把特定类型的数据转换为标准的网络数据。将一些不同关系的或者相同网络中不同的认知片段整合成为一个单一网络, 新组成的单一网络是原来整个网络一些视图或者所有个体网络组合成的整个网络的视图。这个程序输入的文件是一个多元矩阵数据文件, 该文件包含  $K$  个邻接矩阵 (Adjacency Matrices)。每个邻接矩阵的形式为:  $A(i,j)$  嵌入三维矩阵中  $A(i,j,k)$ ,  $k$  表示从  $i$  指向  $j$  关系进行认知或者评价的认知者, 这种形式有益于表示认知社会结构。该程序将三维矩阵变换成二维矩阵的方式有两种, 一种是对  $k$  个矩阵计算所对应的元素总和,  $A^1(i,j)=A(i,j,k)$ , 对  $k$  求总和。该矩阵可以根据一定的临界值等标准进行二值化处理, 从而生成“一致性结构”。另外一种方式是产生一种所谓具有组合性质的“局部汇集结构” (Locally Aggregated Structure, LAS), 方法是让  $A^1(i,j)=A(i,j,i)+A(i,j,j)$ 。这表示在组合的矩阵中, 既定单元格的值是对所涉及的两个行动者的认知的函数 (Krackhardt), 而不是整个网络的。CSS 的目的是根据一定的标准, 将多个具有认知性的社会结构的矩阵组合成为一个矩阵。

## (3) 数据的浏览、展示及描述等操作

这类菜单项比较容易理解。

- **Browse:** 数据矩阵的浏览。选择命令后, 对话框有 File 菜单栏, 通过 File 菜单可以打开本地计算机里存储的数据矩阵; 或者单击窗口右侧的文件夹按钮, 也可以打开数据矩阵文件。确定打开的数据矩阵文件后, 可以通过调整数据显示窗口的工具栏浏览数据。
- **Display:** 在计算机屏幕上展示 Ucinet 数据库。选择数据集名打开后, 程序会产生一个文本文件, 展示该数据集的基本信息, 包括数据集展示的宽度、小数位、行



列值等信息的设定。

- **Describe:** 对数据进行描述（如描述数据的规模、维度、类型及标签等），并可以输入、输出，编辑标签加入到行、列或者整个矩阵当中。

(4) 数据的提取、移动、解压、合并等操作

- **Extract:** 数据提取，即抽取 Ucinet 数据集中的矩阵或指定的行、列数据。使用该功能，首先要选择被抽取的数据矩阵文件名称，数据类型必须是矩阵。主要包括以下 3 种操作（见图 3-4）。

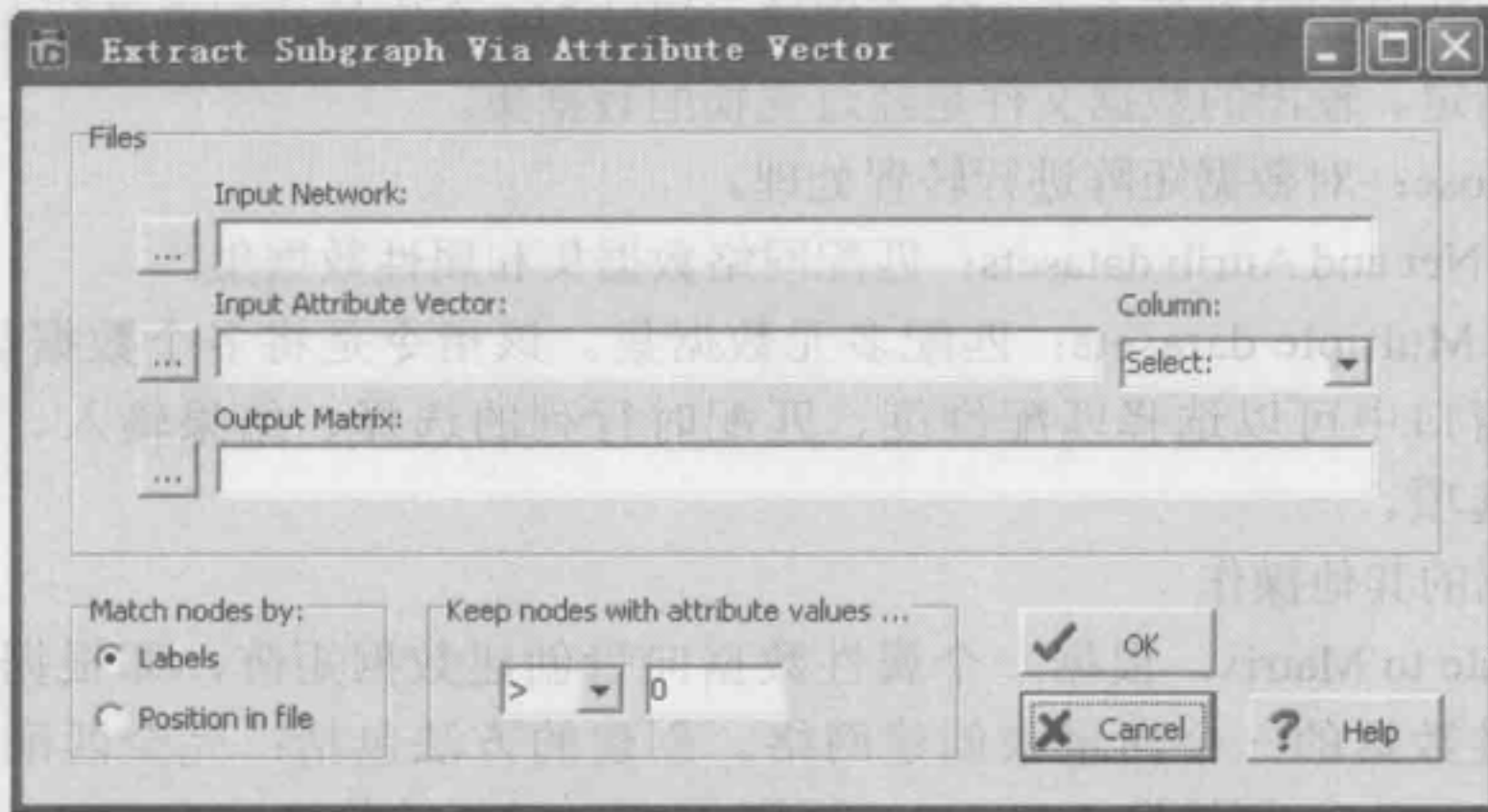


图 3-4 Ucinet 提取功能对话框

➤ **Extract Main Component:** 该指令可以抽取出一个网络的主要弱成分对应的子图。

➤ **Extract Ego-network:** 从整体网络中抽取个体网络 (Egonet)，个体网络是一个点或多个点的个体网邻接矩阵，它包括与一个或多个结点相关的所有关系组成的网络。该指令的对话框需要给出抽取子网络的整体网的文件名，同时需要指定目标结点 (Focal Nodes)，目标结点的指定与 Extract Subgroup 类似，也有一个指定的列表，可以使用关键字 FIRST、LAST、TO 等。是否包括目标结点默认情况下是 Yes，用户可以自行选择目标结点是否包括在新建的个体网中。

➤ **Extract Sub Matrix:** 抽取子矩阵，即从 Ucinet 数据集中抽取部分数据集。抽取同样使用指定列表，过程与上述指令相似，抽取结果是创建了一个带有行列标签的新的数据集。

- **Remove:** 移除 Ucinet 数据库。具体包括两种选项：一种是移除孤立点 (Isolates)，选择该指令，就可以将网络中孤立的点删除，得到关联图；另一种是移除关联小的点 (Pendants)，即与其他成员之间的联系非常少的点。

• **Unpack:** 对一个包含多种关系的矩阵数据进行解压处理，分成多个独立的矩阵数据文件，可根据行、列或整个矩阵进行合并，也可以分解成多个独立的矩阵文件另保存为 Ucinet 数据文件，以便方便进一步的分析。

• **Join:** 该指令的功能是将 Ucinet 的多个数据文件合并成一个数据文件。合并方式可以是根据行、列或整个矩阵进行合并，也可以将多个单独的矩阵或多元关系矩阵合并成一个新的多元关系矩阵。

### (5) 数据的排序、置换、转置等操作

- **Sort:** 按照一定的标准对一个网络中的各个结点进行排序, 以便使其对应所指定的向量秩序。排序分升序和降序, 如果选定升序, 那么向量  $(V_1, V_2, \dots, V_n)$  (每个向量代表一个结点) 会按照从小到大的顺序排列。向量一般可以从 Ucinet 数据矩阵中的行列中选取。
- **Permute:** 按照用户自己指定的顺序列表对数据集的行、列、矩阵进行置换。指定的顺序列表可以是用户自己指定的, 也可以是数据集中的顺序列表。手动键盘输入指定顺序列表必须小于 255 个字符, 超过 255 个字符的列表必须在 Ucinet 数据集中指定。输出的数据文件是经过置换的数据集。
- **Transpose:** 对数据矩阵进行转置处理。
- **Match Net and Attrib datasets:** 匹配网络数据集和属性数据集。
- **Match Multiple datasets:** 匹配多元数据集。该指令是将多个数据集进行匹配, 对话框中可以选择匹配维度、匹配时行列的选择、结果输入、输出时的分类等选项。

### (6) 数据的其他操作

- **Attribute to Matrix:** 根据一个属性数据向量创建数据矩阵, 即根据包含多个点的属性数据的一个向量来创建网络。创建的方法包括: 完全匹配法 (如果向量  $V(i)=V(j)$ , 则矩阵  $X(i,j)=1$ , 否则  $X(i,j)=0$ ); 差值法 (令  $X(i,j)=V(i)-V(j)$ ); 差值的绝对值法; 差值的平方法; 乘积法及数值之和法等。
- **Affiliations (2-mode to 1-mode):** 将 2-模网络 (隶属数据) 转换为 1-模网络数据。分别根据行和列, 可转换成 2 个 1-模网络。即将  $m \times n$  的矩阵通过使用矩阵  $AA^1$  或者矩阵  $A^1A$  两种不同的二进制乘法转换成  $m \times m$  或者  $n \times n$  形式的矩阵。
- **Subgraphs from Partitions:** 根据网络的区分情况抽取子图。
- **Partitions to Sets:** 根据行动者发生矩阵将一个矩阵分区指示向量转换为一个群体, 并根据群体特点展示分区情况。一个分区指标的形式为  $(k_1, k_2, \dots, k_i, \dots)$ , 其中  $k_i$  代表将点  $i$  分配到群体  $k_i$  中。因此分区指示如  $(1\ 1\ 2\ 1\ 2)$  表示点 1、点 2、点 4 被分到第一块; 而点 3、点 5 被分到第二块。一个“点-群体”发生矩阵将点作为矩阵的列, 将群体作为矩阵的行, 如果矩阵中第  $i$  行第  $j$  列的值是 1, 那么代表  $j$  点被分到了  $i$  群体中, 其他情况矩阵中的这些值都是 0。
- **Create Node Sets:** 创建点集, 即在比较两个向量或者一个向量与一个数字的基础上, 创建一个群体指标向量 (Group Indicator Vector)。具体地说, 给定一个包含每个点的属性向量或者是数值向量以及一个临界阈值, 然后指令可以选择出小于 (或者大于) 临界值的行动者。所以临界值本身可以是一个向量, 当行动者的格值小于 (或者大于) 临界值向量的格值时, 该行动者就被选择出来。例如, 两个行动者向量的接近中心度都小于它们的度数中心度, 那么就可以利用上述两个向量比较的方式, 将符合条件的行动者选择出来。这时在指令对话框中需要知道的关系操作符意义如下: LT -Less Than; LE -Less



than or Equal to; EQ -Equal to; NEQ -Not Equal to; GE -Greater than or Equal to; GT -Greater Than.

- Reshape: 重新组织数据, 使之成为规模不同的数据。

### 3. Transform 菜单

Ucinet 转换菜单 (Transform) 包含把一些数据矩阵、图或网络转换成目标类型的命令, 包括以下 3 种类型, 如图 3-5 所示。

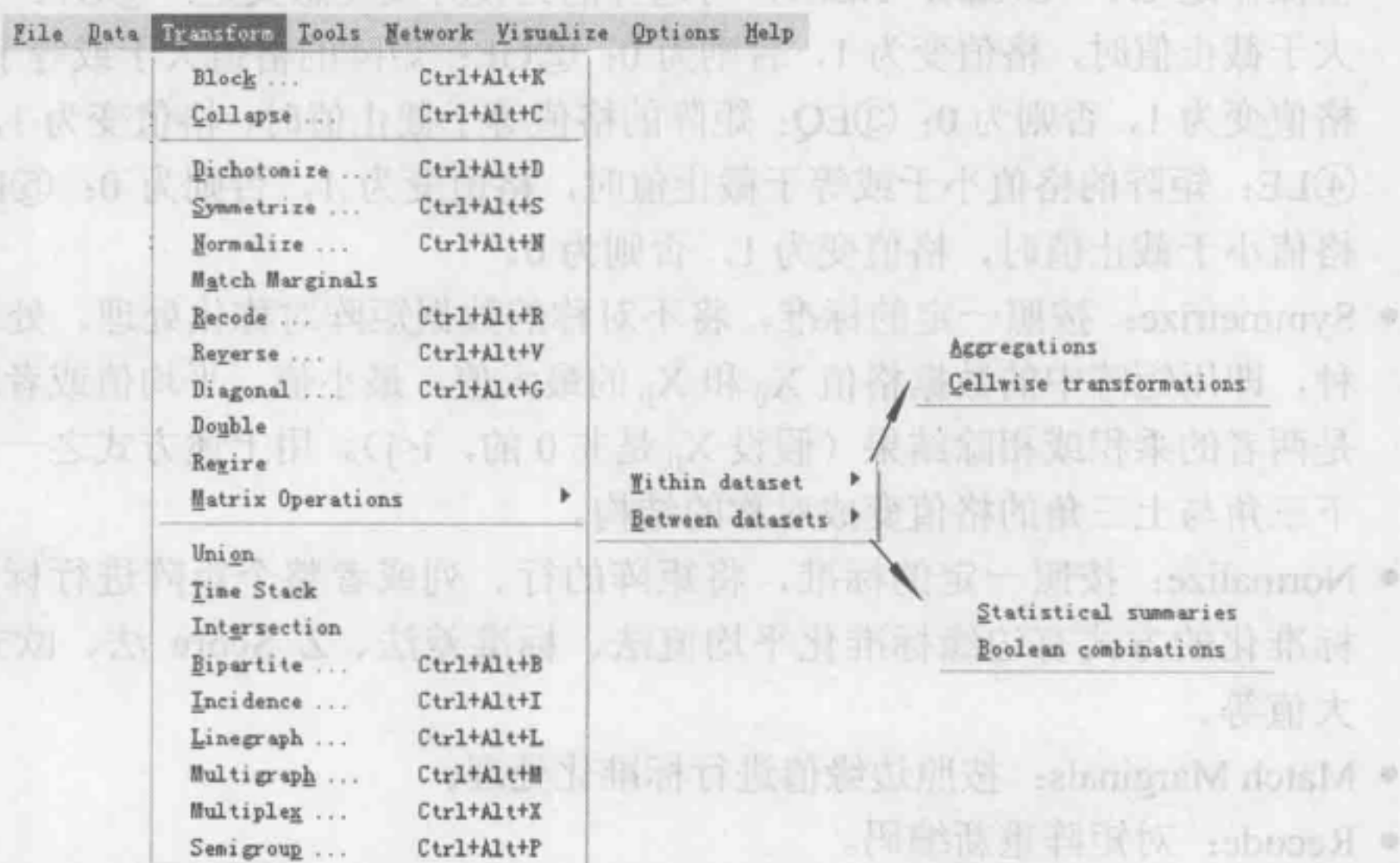


图 3-5 Transform 菜单

#### (1) 第一类命令

第一类命令包括分块 (Block) 和压缩 (Collapse) 两个命令。

- Block: 把一个数据图中的点分块, 并计算块的密度 (Block Densities)、块总和 (Block Sum) 及其他一些统计项。具体地说就是, 邻接矩阵被分成几个子矩阵, 然后计算每个子矩阵的平均值、总和、最大值、最小值、标准差及方差等。这个指令基本上与 Networks | Properties | Density Routine 命令功能相同, 只是 Block 指令提供了计算每个块的更多统计项, 而后者仅仅计算密度。
- Collapse: 压缩, 即把一个矩阵的多行或多列重新组合在一起, 组合成一个更小的矩阵。压缩后的矩阵单元格值可以是原压缩值的平均值、总和、最大值或者最小值。该命令对话框中需要输入压缩指令, 指出哪些行或列需要压缩在一起, 比如需要压缩矩阵的 1、3、4 行成为新的一行, 就需要键入指令: ROWS 1 3 4; 若压缩列 1、2 成为一行, 列 3、4、5 压缩成为一行, 那么指令为:

COLS 1 2

COLS 3 to 5

若需要同时压缩行和列, 需要使用关键字 NODES, 如:

NODES 1 to 3

NODES 4 to 6

## (2) 第二类命令

第二类命令包括 Dichotomize 等 10 个命令，主要针对矩阵的全部格值进行分析，具体分析内容如下：

- **Dichotomize:** 矩阵的二值化处理。该指令是给定一个截止值，然后通过将矩阵的格值与该截止值进行比较进而进行二值化处理。比较的方式有 5 种，默认的截止值操作是 GT (Greater Than)，可选择的比较方式及意义是：①GT：矩阵的格值大于截止值时，格值变为 1，否则为 0；②GE：矩阵的格值大于或等于截止值时，格值变为 1，否则为 0；③EQ：矩阵的格值等于截止值时，格值变为 1，否则为 0；④LE：矩阵的格值小于或等于截止值时，格值变为 1，否则为 0；⑤LT：矩阵的格值小于截止值时，格值变为 1，否则为 0。
- **Symmetrize:** 按照一定的标准，将不对称的数据矩阵对称化处理。处理方式有几种，即用矩阵中的数据格值  $X_{ij}$  和  $X_{ji}$  的最大值、最小值、平均值或者求和，甚至是两者的乘积或相除结果（假设  $X_{ji}$  是非 0 的， $i < j$ ）。用上述方式之一，将矩阵的下三角与上三角的格值变成对称的结构。
- **Normalize:** 按照一定的标准，将矩阵的行、列或者整个矩阵进行标准化处理。标准化的方式有边缘标准化平均值法、标准差法、Z\_Score 法、欧式算法、最大值等。
- **Match Marginals:** 按照边缘值进行标准化处理。
- **Recode:** 对矩阵重新编码。
- **Reverse:** 取相反数，即采用线性转换，将相似性数据转换成距离数据，或者距离数据转换成相似性数据。即矩阵中的最大值和最小值之和减去矩阵中的每个格值。
- **Diagonal:** 对角线命令，可用该命令对方阵对角线上的值进行简单操作。即设置方阵对角线上的格值，并保存。
- **Double:** 将矩阵中的格值加倍。
- **Rewire:** 按照某种标准，对一个矩阵进行最优化处理。可选的标准有 8 种：聚类系数最大化、最小化，距离最大化、最小化，聚类系数最大化并且距离最小化、聚类系数最大化并且距离最大化，聚类系数最小化并且距离最小化，聚类系数最小化并且距离最大化。经过这些标准的处理，矩阵格值会有变化，产生的矩阵成为一个新的满足用户研究的矩阵。
- **Matrix Operations:** 矩阵算法，针对矩阵进行各种计算，包括矩阵内算法和矩阵间算法。其中，矩阵内算法 (Within Dataset) 又包括矩阵内的聚合计算 (Aggregations) 和矩阵内格值的转换 (Cellwise Transformations)。矩阵内聚合计算是求出数据矩阵中行、列、子矩阵或组合的总和、平均值、标准差、最大值、最小值。矩阵内格值的转换是计算矩阵中每个格值的绝对值、对数值、自乘积、负数、平方值等。矩阵间算法 (Between Datasets) 包括统计汇总 (Statistical Summary) 和布尔代数组合 (Boolean Combinations)。统计汇总指令的功能包括求和、最大值、最小值、



平均值、乘积。布尔代数组合指令通过对数据集的条件限制实现矩阵的布尔代数计算。

### (3) 第三类命令

第三类命令包含 Union 等 9 个命令，主要针对矩阵进行其他转换，具体的容如下。

- **Union:** 图的合并。该指令可以选择需要合并的数据集，按一定条件合并图谱。
- **Time Stack:** 将在不同的时间段得到的同一群行动者的关系矩阵合并在一起。
- **Intersection:** 取同一群行动者关系矩阵的交集。
- **Bipartite:** 将一个 2-模数据集矩阵转换成一模的邻接矩阵。任何一个 2-模发生阵都可以看成一个二分图。这个指令将一个 2-模的发生阵转换成一模二分图的邻接矩阵。例如发生矩阵是一个  $n$  行  $m$  列的矩阵，那么转换后的邻接矩阵将会是一个  $(m+n)$  维的方阵。
- **Incidence:** 把一个邻接矩阵转换成发生阵。一个关联矩阵是一个结点的边矩阵。行代表的是结点图，列则代表边。 $i$  行  $j$  列表明结点  $i$  是边  $j$  的事件，这种表示经常被称为超图表示。
- **Linegraph:** 构建线图或者网络。线图  $G$  是通过将  $G$  图的边作为顶点得到的图，若两条边是相邻的，那么两个顶点也会相邻接。
- **Multigraph:** 多图，把一个多值图转换成一组二值图。一个二值图是为多值图的每个不通值而创建的。所创建的二值图都会封装到一个独立的数据集中。二值矩阵中关系的构建取决于多值图中不同值的数量，应该注意不要使用过大的数据集，避免创建出太过大量的二值矩阵。
- **Multiplex:** 从一个多元关系图中构建一个多丛图。
- **Semigroup:** 根据图、有向图以及多元图来构建半群 (Semigroup)。所谓一个网络的半群，指的是组合关系的一种代数表示。比如，一个多元关系中的一系列二值关系矩阵为  $R_1, R_2, \dots, R_n$ ，那么全部可能的一对矩阵的布尔代数乘积给出所有可能的长度为 2 的关系。继续计算长度为 3 的矩阵乘积，直到不出现新的矩阵。以这样布尔乘积的方式构造的全部矩阵就构成了一个半群。该指令可以发现半群的所有成员或一定乘积长度上的半群成员。

## 4. Tools 菜单

Ucinet 工具菜单 (Tools) 包含网络分析者经常使用的技术工具，比如一致性分析、聚类分析等，这些命令又可以分为以下 3 类，如图 3-6 所示。

### (1) 第一类命令

第一类命令包含属性设置、一致性分析、聚类分析、量表及分解。

- **Profit:** 属性设置。
- **Consensus Analysis:** 分析多个量表，分析多个问题回答者在回答方面的一致性。
- **Cluster Analysis:** 对矩阵进行聚类分析，包括层次分析、最适聚类分析、最优化分析。

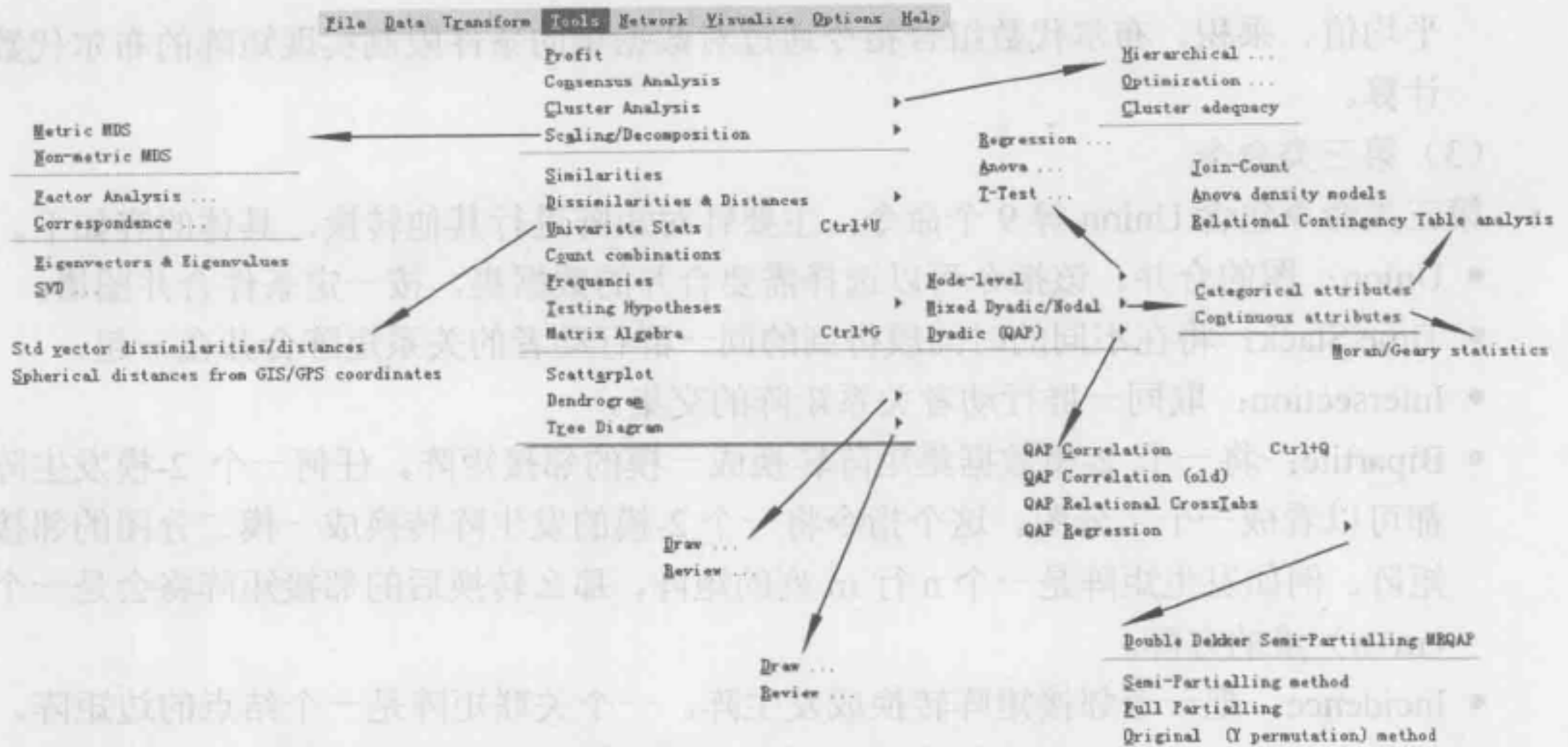


图 3-6 Tools 菜单

- **Scaling/Decomposition:** 量表及分解，它包括以下 3 小类命令：①第一小类，对矩阵进行多维量表分析，包括 Metric MDS 和 Non-metric MDS，即对相近矩阵进行量纲式多维量表分析和非量纲式的多维量表分析。给定一个相近性矩阵（可以是相似性矩阵，也可以是相异性矩阵），该程序会在 k-维空间内发现一系列点，这些点的欧式距离尽可能地接近输入的相似性矩阵；②第二小类，包括因子分析（Factor analysis）和对应分析（Correspondence）。其中，因子分析是指对一个二维矩阵进行完整的因子分析。用主成份方法或最低残差方法对矩阵进行因子分析；对应分析是指对单一的实数矩阵进行对应分析。给定一个非负  $n \times m$  的矩阵 ( $n \geq m$ )，该指令代表着这  $n$  行  $m$  列作为多维空间里常见的向量，该算法对归一化后矩阵执行奇异值分解。矩阵归一化是指调整数据矩阵的行和列，分别归一化以产生更平等的边界；③第三小类，特征向量和特征值分析（Eigenvectors & Eigenvalues）、奇异值分解（SVD），适用于分析 2-模网络数据矩阵。

(2) 第二类命令

第二类命令主要包括以下 7 小类命令。

- **Similarities:** 相似性分析。使用多种算法计算矩阵各行或者各列的相似性系数。
- **Dissimilarities & Distances:** 相异性分析。
- **Univariate Stats:** 对矩阵中的值进行单变量统计分析。该指令对矩阵的行或列、整个矩阵计算平均值、标准差、方差、欧几里得范数、最大值、最小值和总观察值等。
- **Count combinations:** 计算组合频次。
- **Frequencies:** 对行或列进行频次分析。
- **Testing Hypotheses:** 假设检验。它包括以下 3 方面的内容：①Node-Level，点层次的假设检验，本部分包括回归分析、方差分析、T 分布检验；②Mixed Dyadic/Nodal：点-关系混合层次的假设检验，包括对分成两个群组的邻接矩阵进行随机自相关检测、方差密度分析、相关列联表分析等；③Dyadic (QAP)：关系



中其向-关系层次的假设检验,包括矩阵相关分析及旧版的矩阵相关分析、矩阵关系的列联表分析、矩阵回归分析。

- **Matrix Algebra:** 矩阵代数分析。该指令可执行矩阵的乘积、逆矩阵、特征值等运算。

(3) 第三类命令。

第三类命令包括以下3个画图命令。

- **Scatterplot:** 散点图。
- **Dendrogram:** 树形图。
- **Tree Diagram:** 树图。

5. Network 菜单

Ucinet 网络菜单 (Network) 包含网络分析者经常使用的技术工具,比如一致性分析、聚类分析等,这些命令又可以分为以下4类,如图3-7所示。

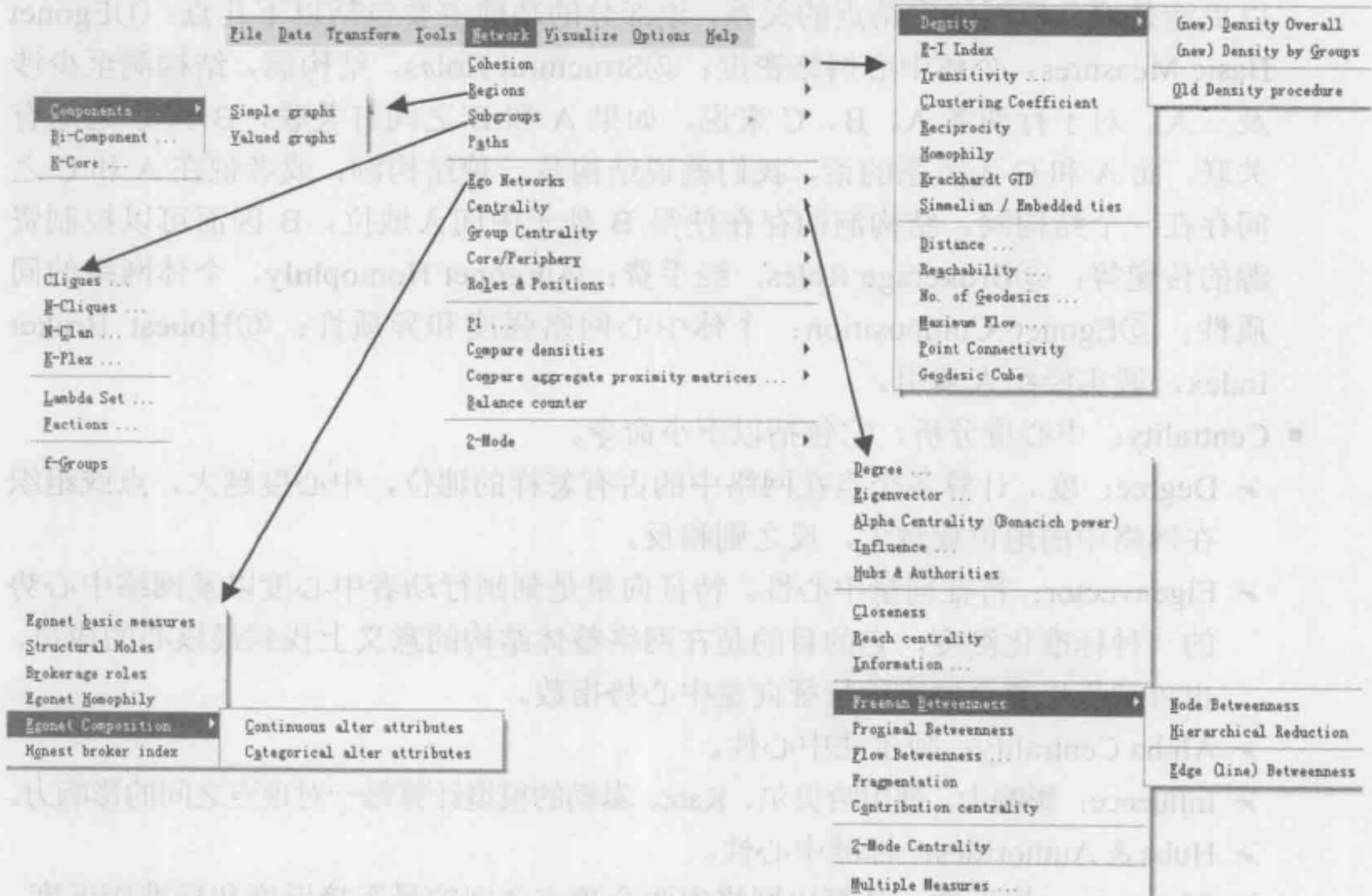


图3-7 Network 菜单

(1) 第一类命令

第一类命令包括以下4个命令。

- **Cohesion:** 凝聚子群分析,主要分析网络关系的紧密程度。该类功能包括较多的小指令,分别是: ①Density, 密度分析,包括整体网络的密度分析、网络群组密度的分析、旧版密度分析; ②E-I Index, 计算网络某一个部分的 E-I 指数,执行变换测试来评估它的意义; ③Transitivity, 测定网络中具有传递关系的三元组的密度; ④Clustering Coefficient, 计算每个行动者的聚类系数及整个网络的聚类系数和加权聚类系数; ⑤Reciprocity, 互惠基础上的凝聚子群分析。建立在互惠基

基础上的凝聚子群多指派系，派系是最大的完全子图，包含三点以上，不能向其中加入新的点。即在该点集中，任何点对之间都存在一条直接相连的线，并且该派系不能被其他任何派系所包含；⑥Homophily，网络同质性分析；⑦Krackhardt GTD，社会网络的内聚测量，利用连通性、层级、效能和LUB等指标对矩阵内聚力进行测量分析；⑧Simmelian/Embedded ties，连带关系分析。

- **Regions**: 区域分析，该功能包括成分分析 (Components)、二分成分分析 (Bi-Components)、K-核分析 (K-Core)。
- **Subgroups**: 群组分析。包括派系分析、N 派系分析、N-宗派分析、K-丛分析等。
- **Paths**: 路径。寻找不同结点之间的路径。

## (2) 第二类命令

第二类命令包括以下 5 个命令。

- **Ego Network**: 个体中心网络。个体网络是从整体网络中抽取出的一个子图，可以更清楚地分析网络中结点的关系，该部分的功能主要包括以下几点：①Egonet Basic Measures，个体中心网络密度；②Structural Holes，结构洞，结构洞至少涉及三人。对于行动者 A、B、C 来说，如果 A 和 B 之间有关联，B 和 C 之间有关联，而 A 和 C 无关系的话，我们就说结构是一种结构洞，或者说在 A 和 C 之间存在一个结构洞。结构洞的存在使得 B 处于中间人地位，B 因而可以控制资源的传递等；③Brokerage Roles，经手费；④Egonet Homophily，个体网络的同质性；⑤Egonet Composition：个体中心网络强度和异质性；⑥Honest Broker Index，诚实经纪人索引。
- **Centrality**: 中心度分析，它包括以下小命令。
  - **Degree**: 度。计算各个点在网络中的占有怎样的地位，中心度越大，点或组织在网络中的地位就越大，反之则相反。
  - **Eigenvector**: 特征向量中心性。特征向量是刻画行动者中心度以及网络中心势的一种标准化测度，它的目的是在网络整体结构的意义上找到最核心的成员，也可测量出整个网络的特征向量中心势指数。
  - **Alpha Centrality**: 阿尔法中心性。
  - **Influence**: 影响力。使用哈贝尔、Katz、泰勒的模型计算每一对顶点之间的影响力。
  - **Hubs & Authorities**: 标准中心性。
  - **Closeness**: 接近性。计算出网络中两个顶点之间的最远接近度和标准接近度，在此基础上给出整个网络的接近中心度。如果一个点与网络中所有其他点的距离都很短，则称该点具有较高的整体中心度（又称接近中心度）。
  - **Reach centrality**: 到达中心度。计算每个结点的结点数量是否达到 k。对 k=1，这相当于度数中心性。在有向网络中，接入度数和发出度数都被算在计算范围内。
  - **Information**: 信息中心度。为每个顶点计算斯蒂芬森和 Zelen 信息中心度，然后给出一个整体网络信息的集中指数。
  - **Freeman Betweenness**: 自由中间度。
  - **Proximal Betweenness**: 最接近中间度。
  - **Flow Betweenness**: 流中间度。计算每个顶点的流中间状态和归一化流中间状



态的向心性，并给出整个网络的中间性中间度。

- Fragmentation: 总体分裂性。
- Contribution Centrality: 贡献中心度。
- 2-Mode Centrality: 2-模网络中心性。
- Multiple Measures: 多重方式。
- Group Centrality: 群组中心性。包括寻找特定规模的具有最高中心度的群组(Find)和执行测试(Test)来评估是否指定的群组有一个较高的中心分数。
- Core/Periphery: 核/外围。使用遗传算法等识别哪些结点属于核心结点，哪些结点属于边缘结点。
- Roles & Positions: 角色&位置，包括以下几个方面：①Structural，结构，包括轮廓、Concor、优化；②Automorphic，自同构，包括 MaxSim、全部排列(All Permutations)；③Exact，精确的优化；④Maximal Regular，最大规则。

### (3) 第三类命令

第三类命令包括以下4个命令。

- P1: 使二进制网络符合荷兰和 Leinhardt P1 模型。
- Compare densities: 比较密度。比较两个网络的密度大小。
- Compare aggregate proximity matrices: 比较合计邻接矩阵。
- Balance counter: 平衡计算器。

### (4) 第四类命令

第四类命令包括一个命令，即 2-Mode: 使用遗传算法对 2-模网络进行变换运算。

## 6. Visualize 菜单

Ucinet 可视化菜单 (Visualize) 主要是利用软件对 Ucinet 处理后的数据进行可视化分析。可视化软件主要有 NetDraw、Mage 和 Pajek，如图 3-8 所示。选择其中的一个命令，会调出相应的画图软件，对数据矩阵进行可视化处理。

## 7. Options 菜单

Ucinet 选项菜单 (Options) 可以通过提供的一些命令改变 Ucinet 系统参数的默认设置值，包括以下菜单选项，如图 3-9 所示。

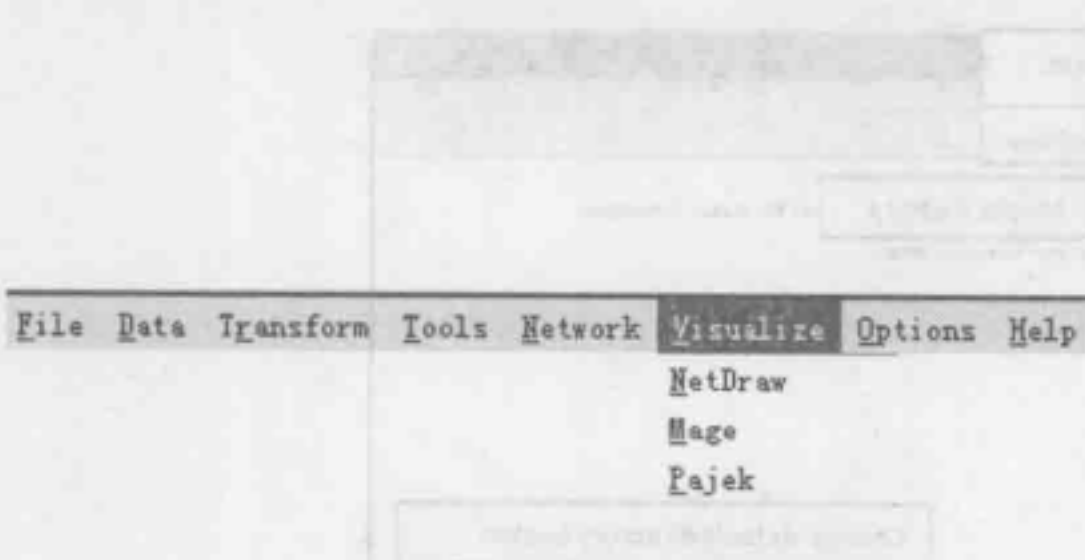


图 3-8 Visualize 菜单

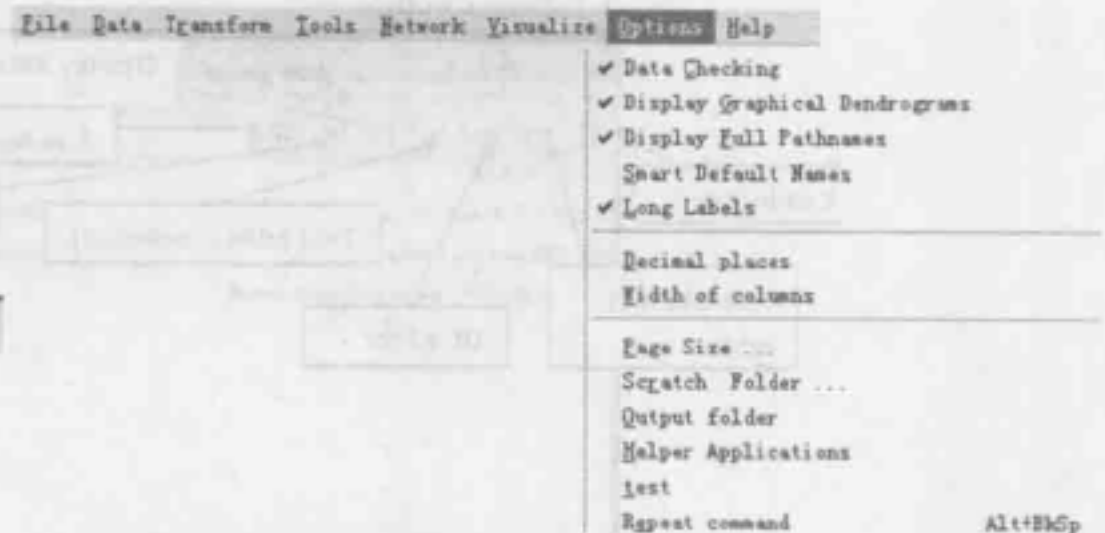


图 3-9 Options 菜单

- Data Checking: 数据检查。
- Display Graphical Dendrograms: 展示图形。
- Display Full Pathnames: 展示完整的路径名。

- Smart Default Names: 优化默认名。
- Long Labels: 设定使用长标签
- Decimal places: 设定小数点的位置, 一般确定到小数点的后 3 位。
- Width of columns: 设定列的宽度。
- Page Size: 改变页码的大小。
- Scratch Folder: 撤销文件夹。
- Output folder: 输出数据所在的文件夹, 可以修改。
- Helper Applications: 应用帮助命令。
- test: 计算数据行或列的不相似性。
- Repeat command: 重复刚刚使用的命令。

### 8. Help 菜单

帮助菜单 (Help) 包含注册信息、帮助标题、技术支持、关于等信息, 如图 3-10 所示。

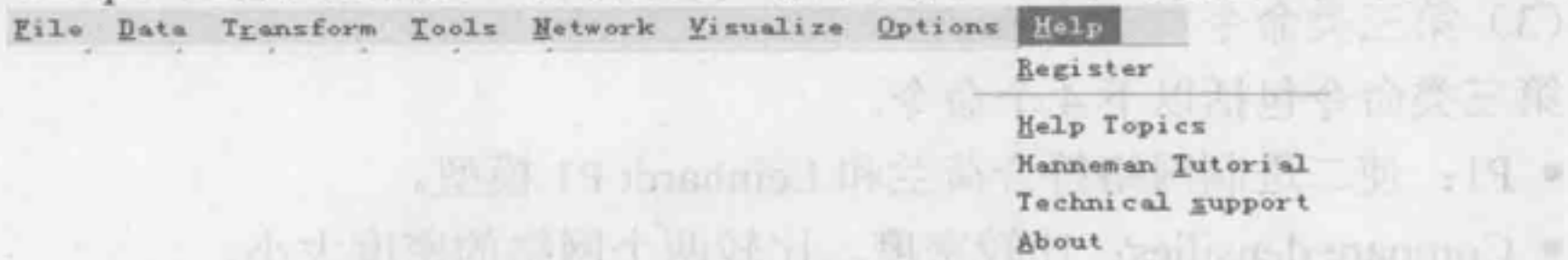


图 3-10 Help 菜单

### 3.2.2 工作界面

用户可以进入 Ucinet 网站 (URL: <http://www.analytictech.com/Ucinet/description.htm>), 选择 Downloading 进行下载。可以将文件 Ucinetsetup.exe 保存在本地机器中的文件夹中, 双击该文件, 安装程序便开始运行, 使用者可以在提示的引导下完成安装 (一般用户可以完全按默认设置安装)。值得注意的是, 安装过程中会更新软件数据包中自带的那部分数据集, 但不会对用户自己的数据进行更改。

启动 Ucinet 以后, 将出现 Ucinet 主界面, 如图 3-11 所示。

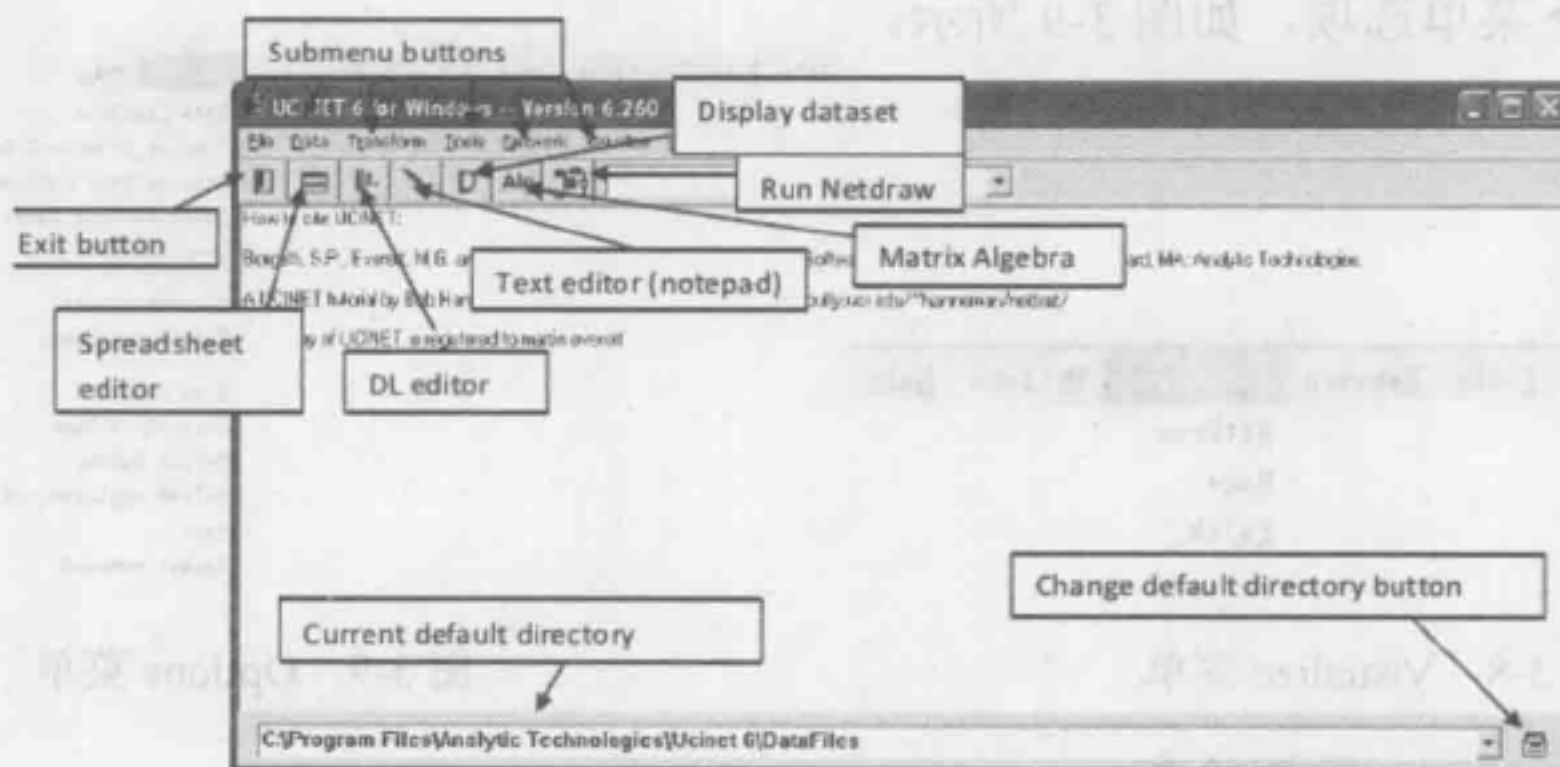


图 3-11 Ucinet 主界面

菜单栏涉及了 Ucinet 的所有功能, 分为文件、数据、转换、工具、网络、可视化、



选项和帮助。位于菜单栏下方的按钮是常用菜单的快捷方式，共有7个，从左到右分别是退出、数据表编辑器、DL 数据语言编辑器、文本编辑器、显示数据库、矩阵算法、运行 Netdraw 软件。选择相应菜单或按钮，窗口会自动在下方运行；界面最底端地址栏显示当前数据的默认存储目录，如需更改可单击右侧按钮更改默认目录。

运行 Ucinet 程序，通常需要指定一个 Ucinet 数据集，给出一些参数。Ucinet 自身可能自带一些默认的参数，这些是可以根据需要修改的。Ucinet 伴随大量的标准数据集，这些标准数据集都放置在默认目录下。每当运行 Ucinet 的一个程序，屏幕上就会出现关于结果的文本文件，通常这些结果存储在 Ucinet 默认目录下的 data 文件夹中。

下面，我们运行 Ucinet 进行分析，data 计算在一个称为 TARO 的标准 Ucinet 数据集的全体参与者的权重。首先选择 Network | Centrality | Degree 命令，如图 3-12 所示。

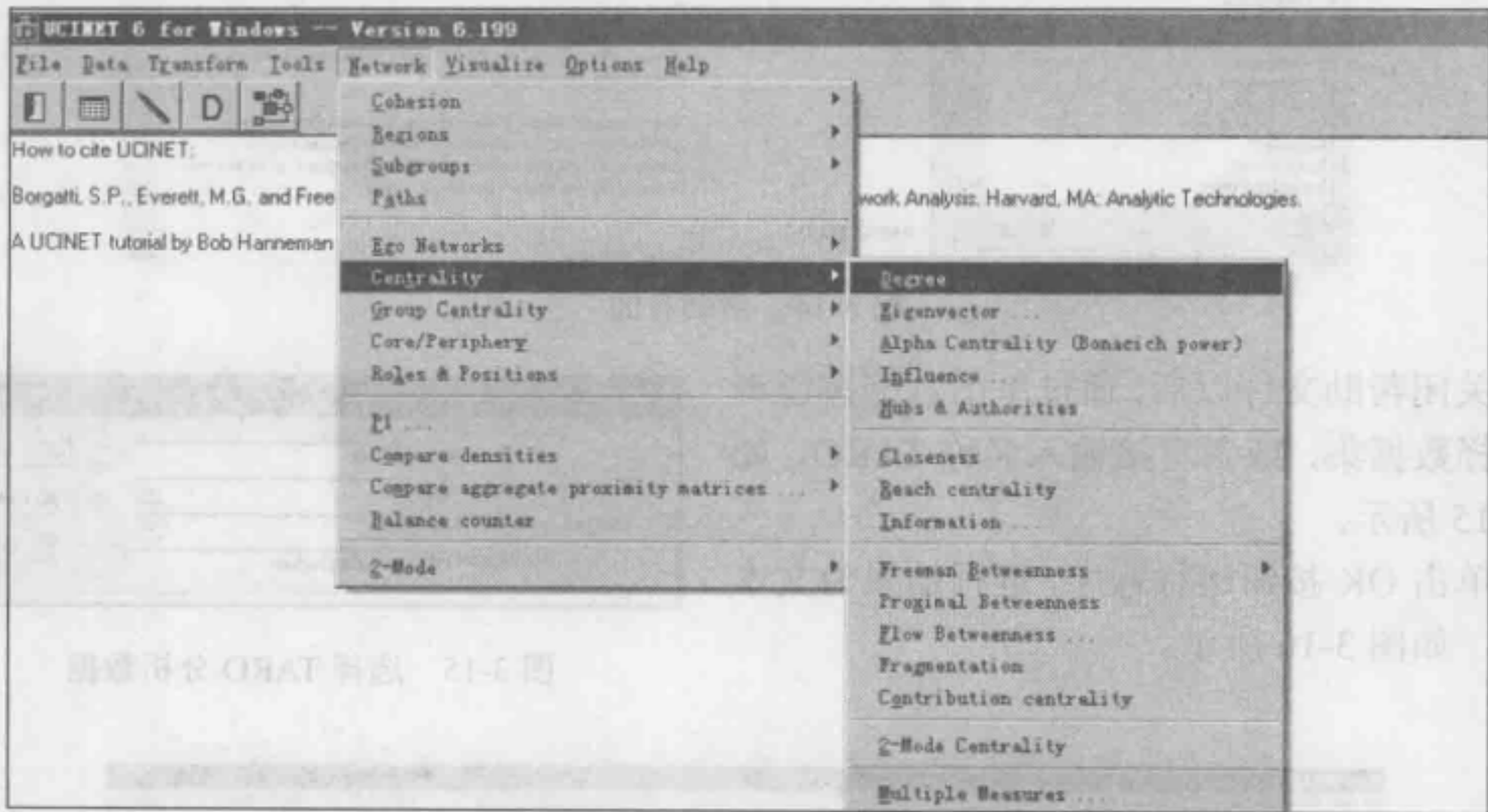


图 3-12 选择 Network | Centrality | Degree 命令

将会显现如图 3-13 所示的对话框：选择数据集及参数，单击 OK 按钮，程序会自动运行。

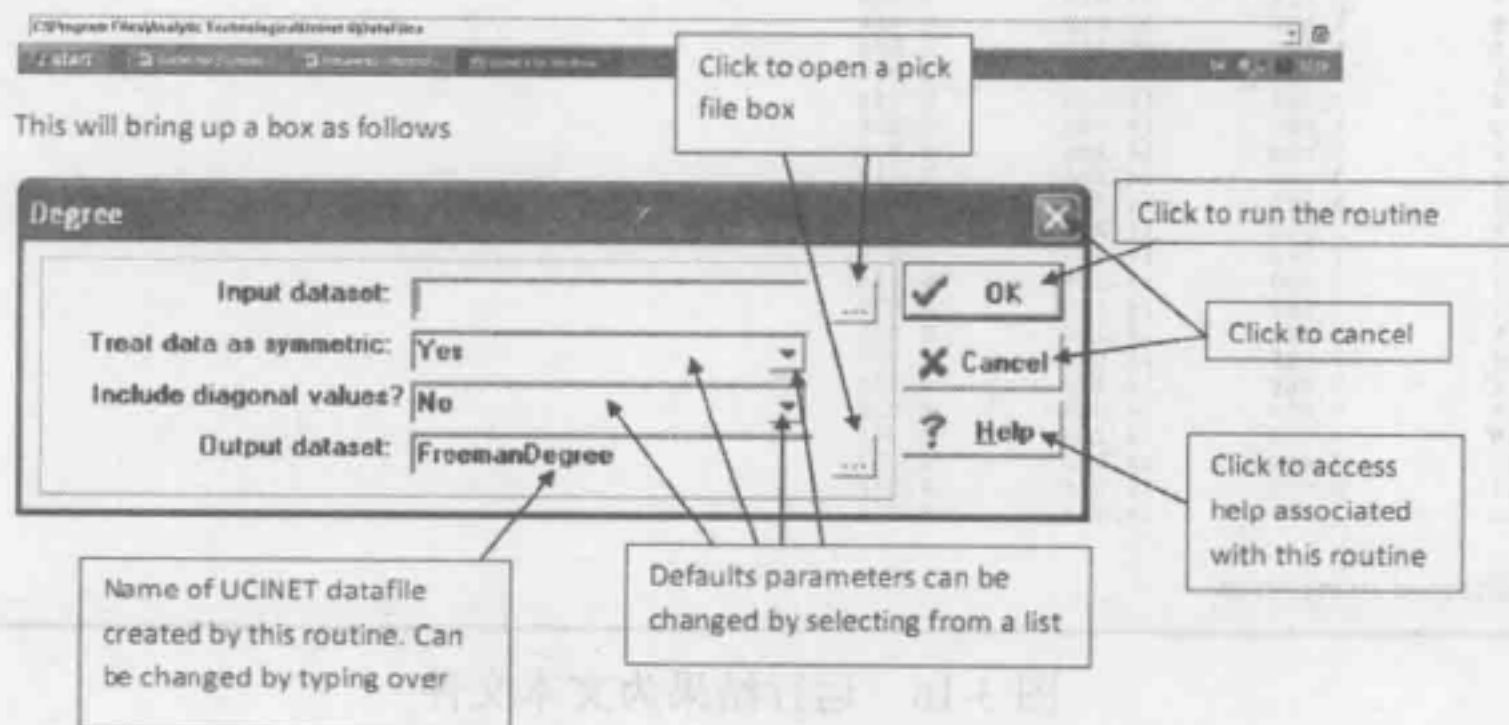


图 3-13 单击后显现的对话框

若单击了 Help 按钮，帮助界面就会在屏幕上打开，如图 3-14 所示。帮助文件给出

了一个程序的详细介绍，会解释参数并描述在记录文件和屏幕上显示出来的输出信息。

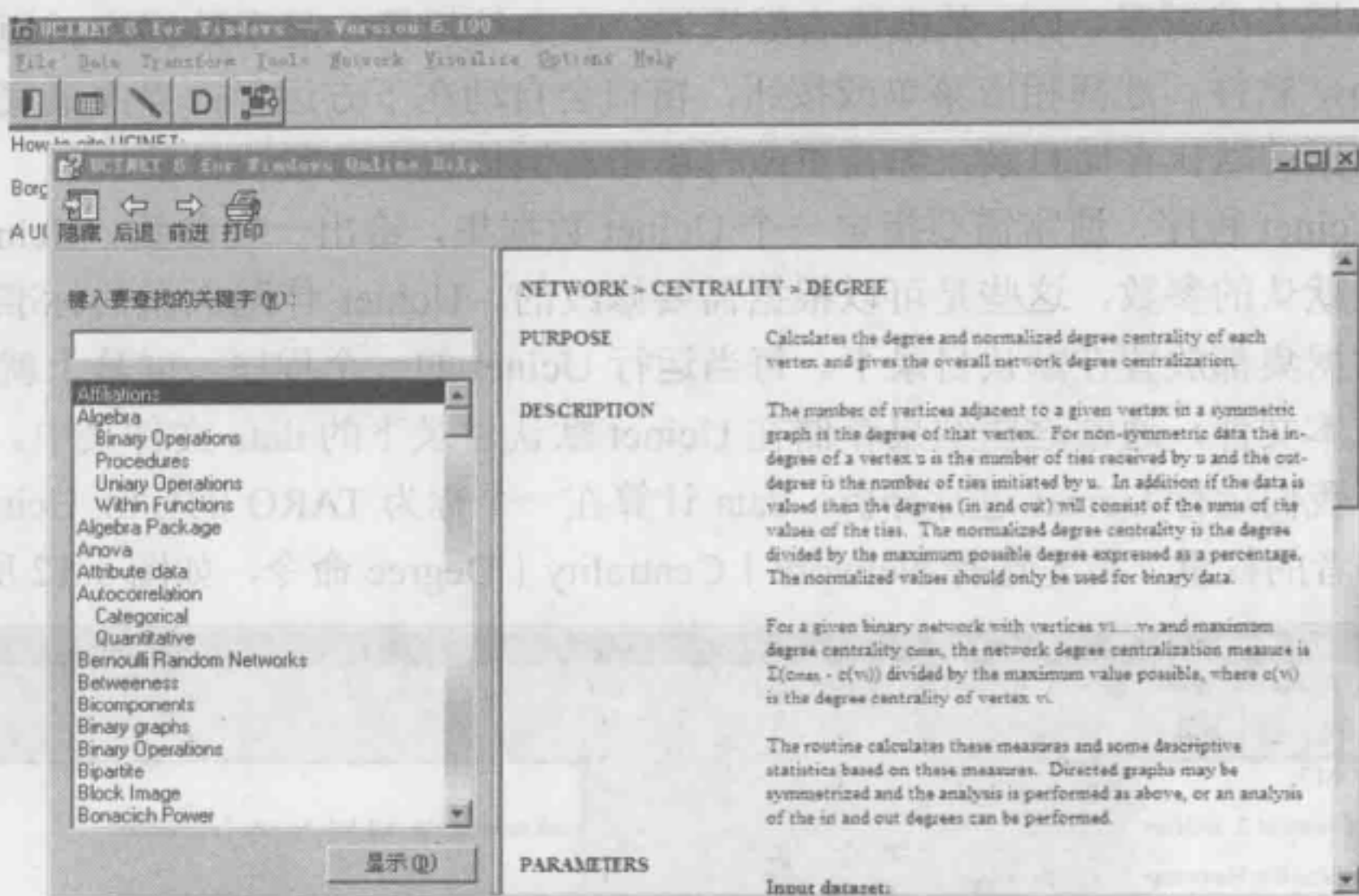


图 3-14 帮助界面

关闭帮助文件以后，通过单击右侧按钮浏览选择数据集，或者直接输入名称 TARO，如图 3-15 所示。

单击 OK 按钮运行程序，运行结果为文本文件，如图 3-16 所示。

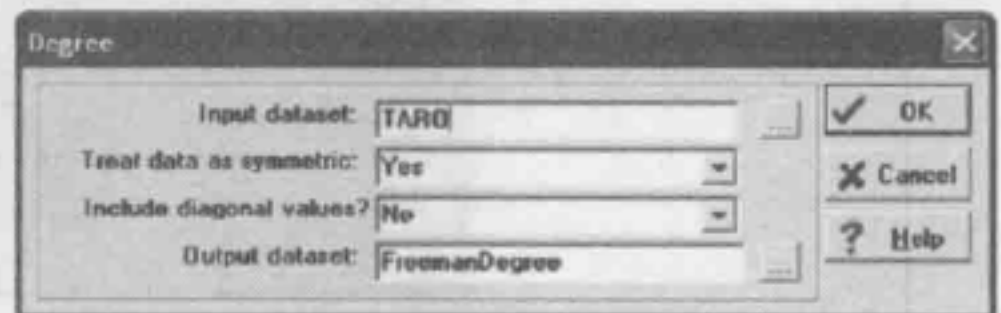


图 3-15 选择 TARO 分析数据



图 3-16 运行结果为文本文件

这个文件可以保存或复制、粘贴到 Word 文档中，当 Ucinet 被关闭时，这个文件将会被删除。需要指出的是，当这个程序运行时，也创建了一个名为 Freeman Degree 的新 Ucinet 文档。我们可以使用 Data|Display（快捷键为【Ctrl+D】）查看新的 Ucinet 文件，



或者选择单击工具栏中的 D 按钮。单击后弹出 Display datasets 的对话框，在对话框的 Data Set Filename 选框中填写需要打开的文件名 Freeman Degree 或者直接单击右侧的选择按钮浏览选择该文件，最后单击 OK 即可浏览文件内容，得到如图 3-17 所示的文件。



图 3-17 选择 Freeman Degree 后得到的文件

使用电子表格编辑器可以修改任何数据或输入新的数据，这对于传输 Ucinet 数据（如中心得分）到 Microsoft Excel 或 SPSS 也是非常有用的。另外，DL 格式提供了一种输入数据时更复杂的灵活方式——电子表格编辑（Spreadsheet Editor），该功能与各种统计分析的运算方法结合，可以与多种软件进行数据交换。

打开电子表格编辑器有两种途径：选择 Data | Spreadsheets | Matrix 命令（快捷键为【Ctrl+S】）；或是直接单击快捷菜单中第二个快捷按钮，打开电子表格编辑器，将获得如图 3-18 所示的界面。根据界面重要按钮提示，用户可以轻松地输入数据并加以保存，以供下一步分析使用。

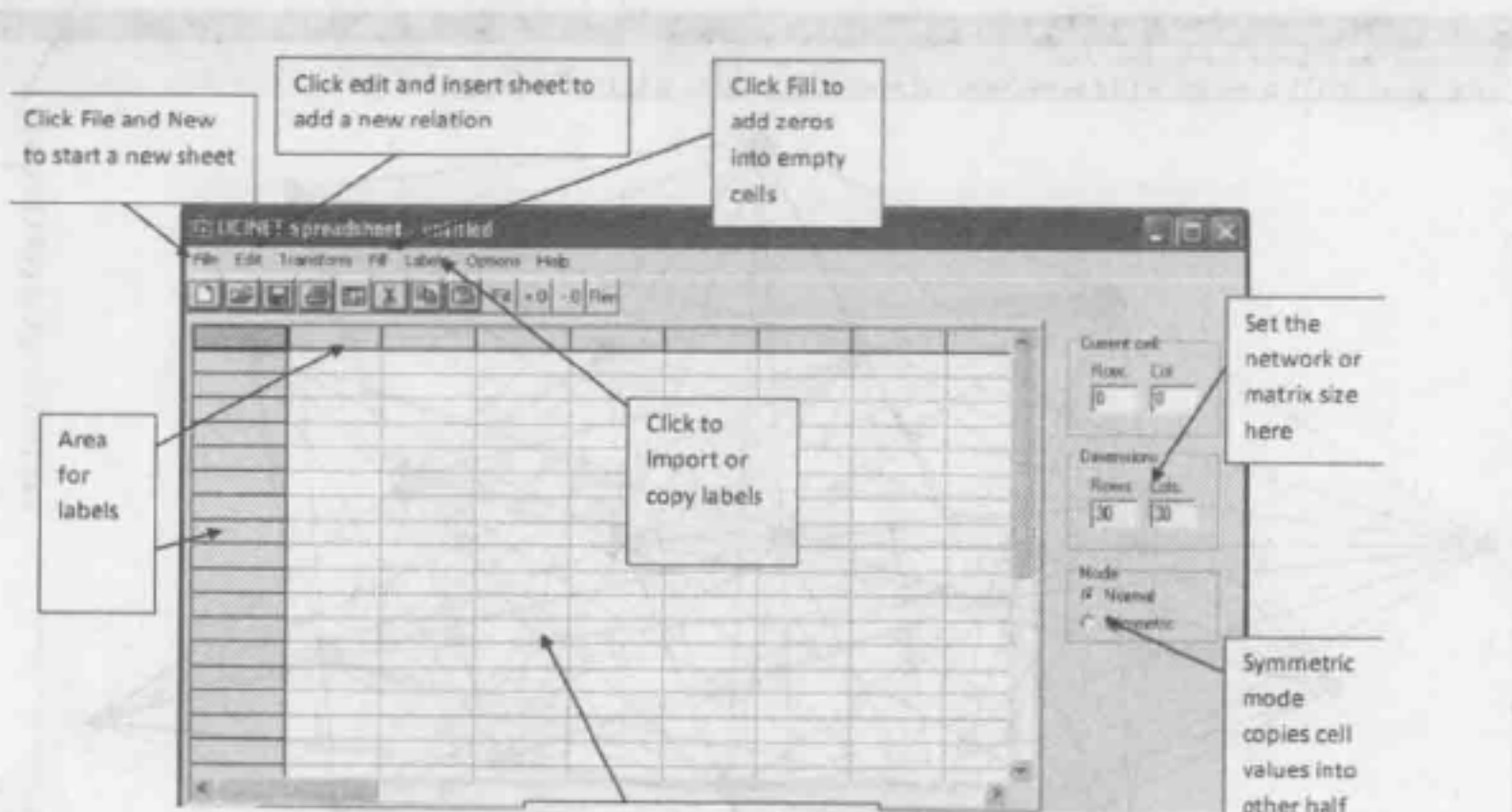


图 3-18 打开电子表格编辑器界面

Ucinet 还可以使用 Netdraw 进行可视化分析。选择 Visualize | Netdraw 命令，打开可视化界面，如图 3-19 所示。

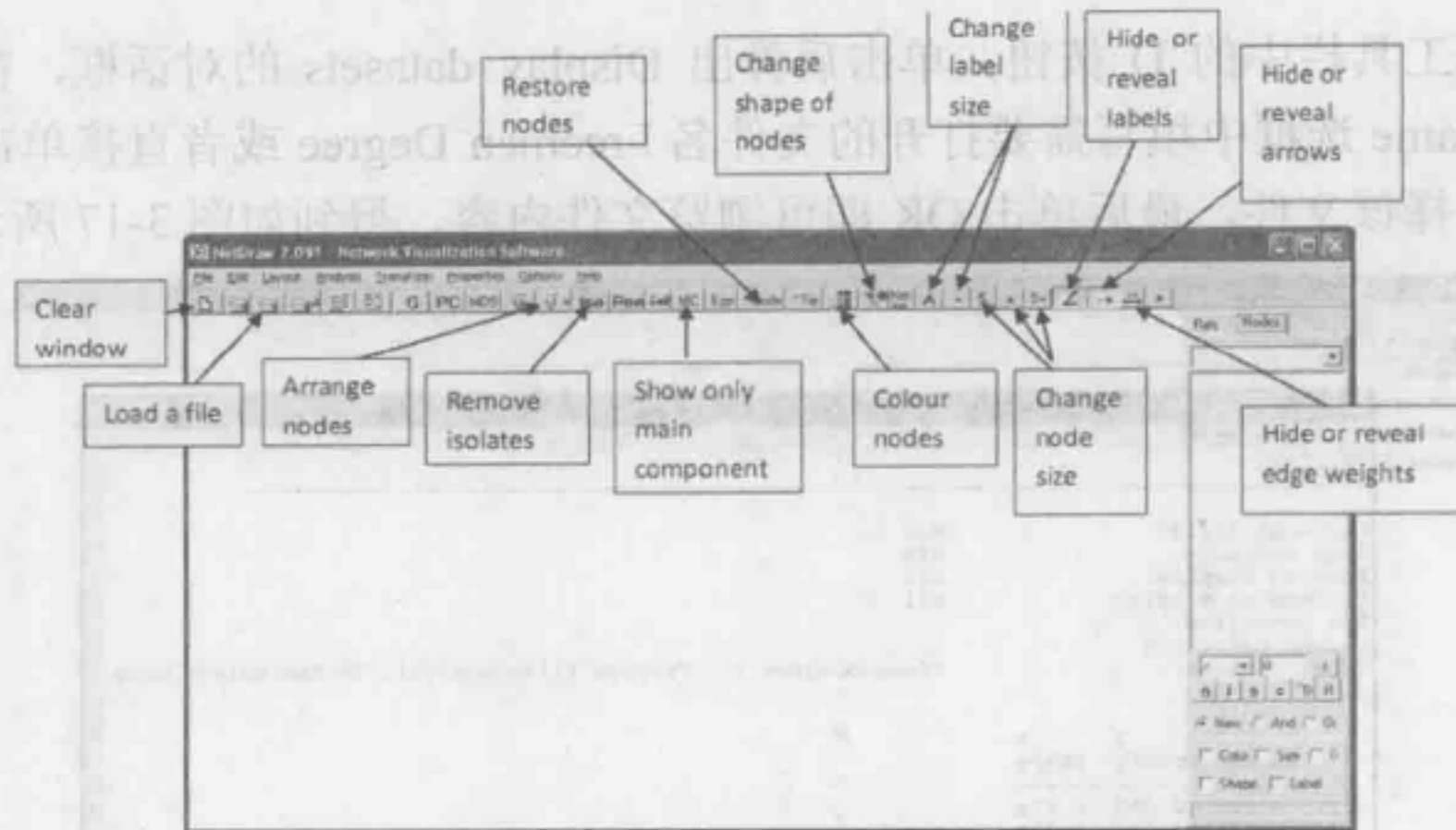


图 3-19 可视化界面

选择 File | Open 命令，在弹出的对话框中选择软件自带数据集 Krack-High-Tec.###h，如图 3-20 所示。

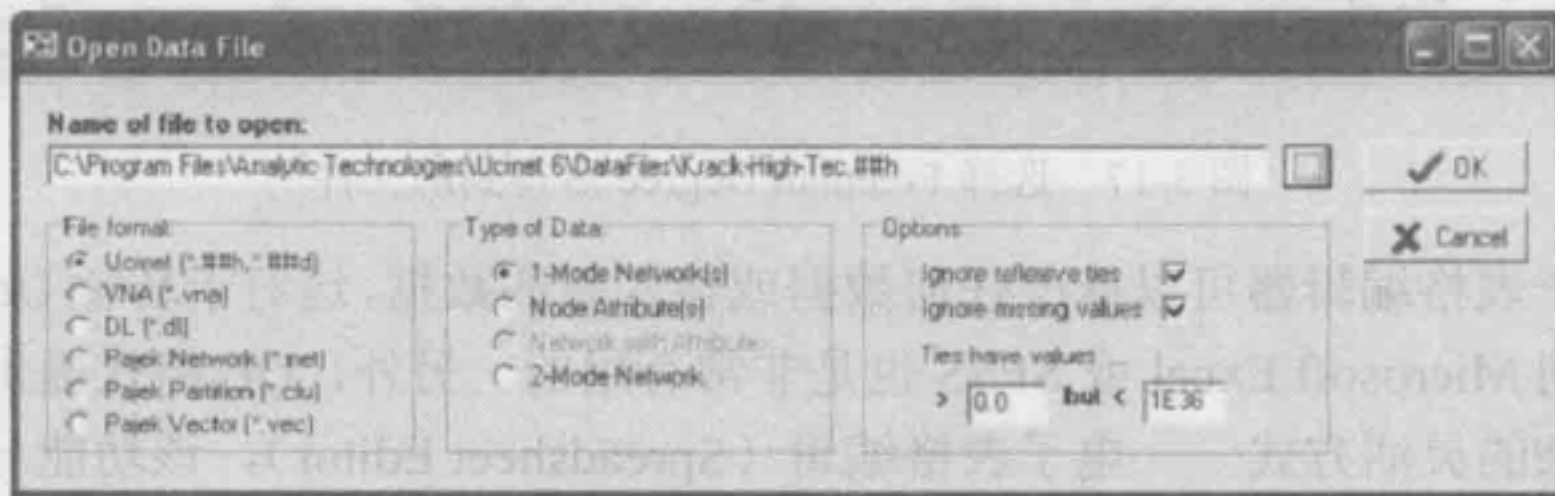


图 3-20 选择软件自带数据集 Krack-High-Tec.###h

然后单击“确定”按钮，应该看到如图 3-21 所示的界面。单击显示在这里的 Rels 标签，将看到这些数据有 3 个相关 Advice、Friendship and Reports to。如果一种相关被勾选，那么关系到它的边会被显示。

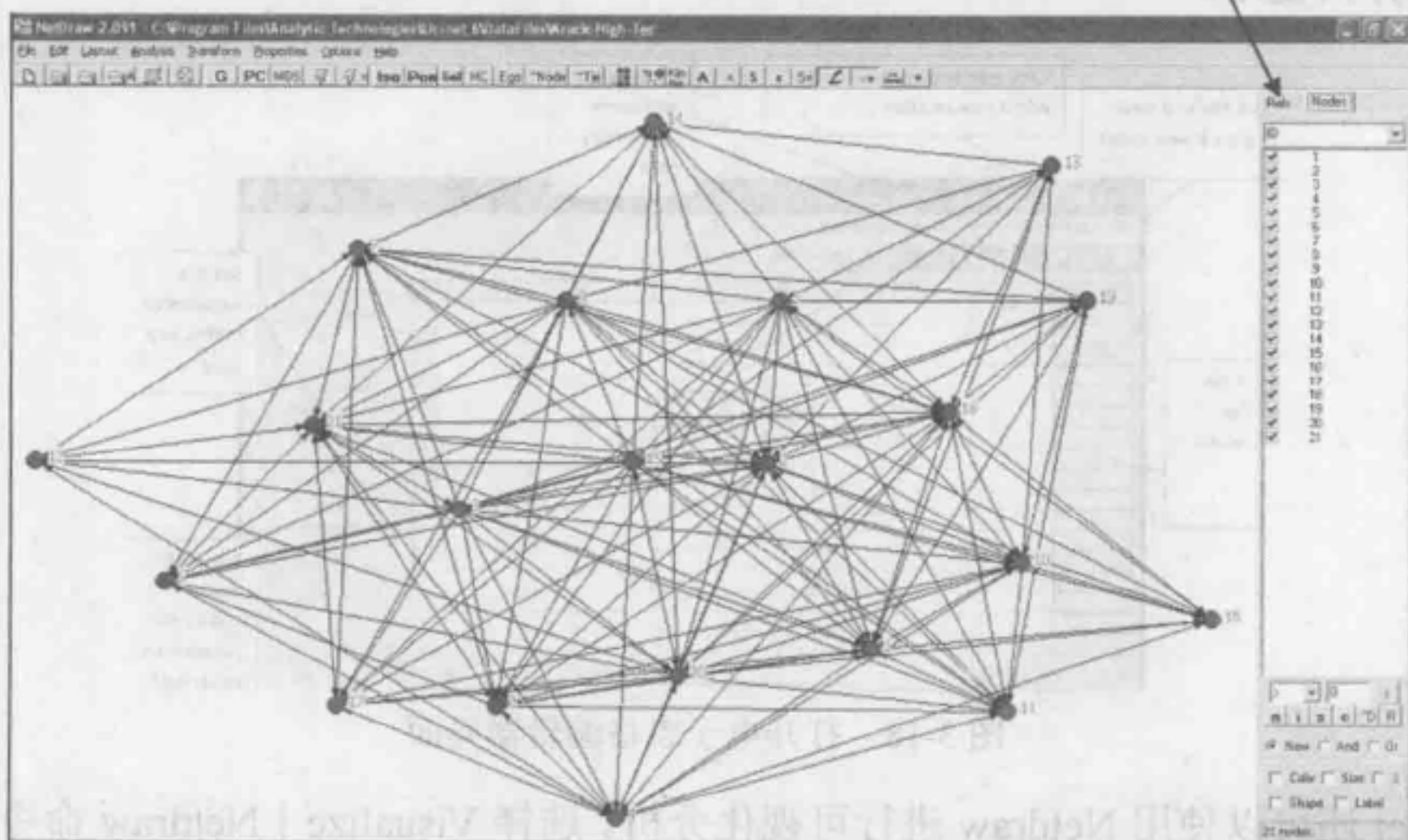


图 3-21 单击确定按钮后看到的界面



以上介绍的是初步使用 Netdraw 进行数据集分析的步骤。使用 Netdraw 进行分析时，还可以通过引进数据集的属性文件、选择不同算法、改变结点大小颜色等来进行可视化分析。

### 3.3 Ucinet 数据准备

#### 3.3.1 数据格式及输入

在使用 Ucinet 6.0 进行任何分析之前，必须创建一个 Ucinet 数据集。通常情况下，网络数据的获取是通过问卷的形式或是从书籍和期刊上获取的数据表条目（如各国之间进行贸易的数据）。在这两种情况下，数据的载体都是纸质形式的，所以必须将纸质载体上的数据输入到计算机的文件中。最常见的方法是使用一个文件编辑器或是你最喜欢的文字处理器将数据输入到一个 ASCII 文件中。输入的数据可以是本章中介绍的任何一种数据格式。一旦数据记录在计算机磁盘上，就可以使用数据导入/导出命令将这些数据转换成一个 Ucinet 6.0 数据集。对于用户统计软件比如 SYSTAT 软件，这应该是一个常见的步骤。在 SYSTAT 软件中，使用数据命令读取一个 ASCII 文件并创建一个 SYSTAT 系统文件，然后软件的统计程序就可以读取了。SPSS 和 SAS 也有类似的统计分析指令，只是该软件不需要创建永久的系统文件。

Import 指令可以处理各种文件类型上的 ASCII 数据。一般来说，支持 4 种数据文件：初始数据（Raw）、数据语言数据（DL）、Excel 和 Ucinet 3（Ucinet 6.0 与 Ucinet 5.0 使用相同的数据集，这些不能被导入）。初始数据（Raw）只包含数字，如一个被“调查者-变量”矩阵记录的是回答者在回答问题时得到的数值型编码数据。DL 文件包含数据与初始数据一样，但还包含关键字及描述数据的信息，比如文件类型、行和列的数量、变量名、该文件的名称等。Ucinet 3 文件类似于 DL 文件，但对数据信息的提供要求更严格。Excel 是标准的 Excel 电子表格数据文件。

如果是从头开始输入数据，建议使用 DL 文件类型（可以在任何时间导出数据到另一个文件类型）。这个文件类型是最灵活的，能接收各种格式的数据。

无论导入的文件类型是哪一种，结果都是一样的：一个 Ucinet 6.0 数据集，可以作为任何数值程序的输入。然而应该注意，可以作为数据的存储方式仅有几个选择。Ucinet 有 3 种基本数据类型：字节型、短整型、实数型。当从一个 ASCII 文件导入数据，会被要求选择这 3 种类型之一。除非是一个非常大的数据集，最好的建议是保留默认设置类型为实数型。

- 实数型数据类型是最强大的。实数型数据集可以包含值的范围从  $-1E36 \sim +1E36$ 。它们也含有缺失值，缺失值的内部存储值为  $1E38$ 。实数型数据类型的缺点是：4 B 的存储必须代表每个值，这意味着相当庞大的文件。例如，一个实数型数据集，其中包含两个  $150 \times 150$  的矩阵就需要 176 KB 的磁盘空间。
- 短整型（Smallint）数据类型每个值只需要 2 B 的存储，但只能代表整数范围在  $-32000 \sim +32000$ 。缺失值是不允许的。



- 字节型数据是最节省内存的，每个值只需要 1 B 的存储，通常是二进制数据的最佳选择。字节的数据集可以包含的整数范围从 0~255，不允许缺失值。

如果需要缺失值，必须使用实数型数据集。然而，大多数网络分析不允许缺失值，所以只有少数 Ucinet 程序知道如何处理它们。那些不能处理缺失值的程序会将缺失值自动转换成 0 或其他合理的值。

在选择一个数据类型时，应该意识到，选择不同的数据类型影响着存储数据的磁盘空间，但它没有影响程序需要处理数据的内存空间。例如，MDS 程序是处理实数型数据文件的，无论数据在磁盘上是如何存储的，程序读取的时候都是实数型数据。同样，派系 (Clique) 等只能处理整数型数据的程序，将自动对实数型数据取整，将它们读取到内存中。

对于缺失值要注意，所有大于 1E37 的数据值被认为是缺失的。在一个 ASCII 文件中，发现的任何非数字字符也被认为是缺失值。例如，下面的矩阵包含 3 个缺失值：

```

0      9.7    1E38    .
16.    0      3.1     na
18.1   1e9    0      -1.2
a3     12     .013    0

```

注意：Ucinet 6.0 不像 SYSTAT 软件，读取一个句点时不是缺失值。如果导入一个 SYSTAT 的数据集，其中包含缺少的值，应该使用一个文本编辑器将所有句点为 NA。

下面介绍一下 Ucinet 常用的 3 种数据格式。

### (1) 第一种数据格式

第一种数据格式是初始数据文件 (Raw Data File)，仅仅包含数字，不包含数据的行数、列数、标签、标题等信息，因而只能以矩阵的形式输入。在 Ucinet 6.0 中，选择 File | Text Editor 命令或者直接单击主界面的第四个快捷按钮，即可以打开一个文本编辑器，用户可以直接在里面输入初始数据，如图 3-22 所示。

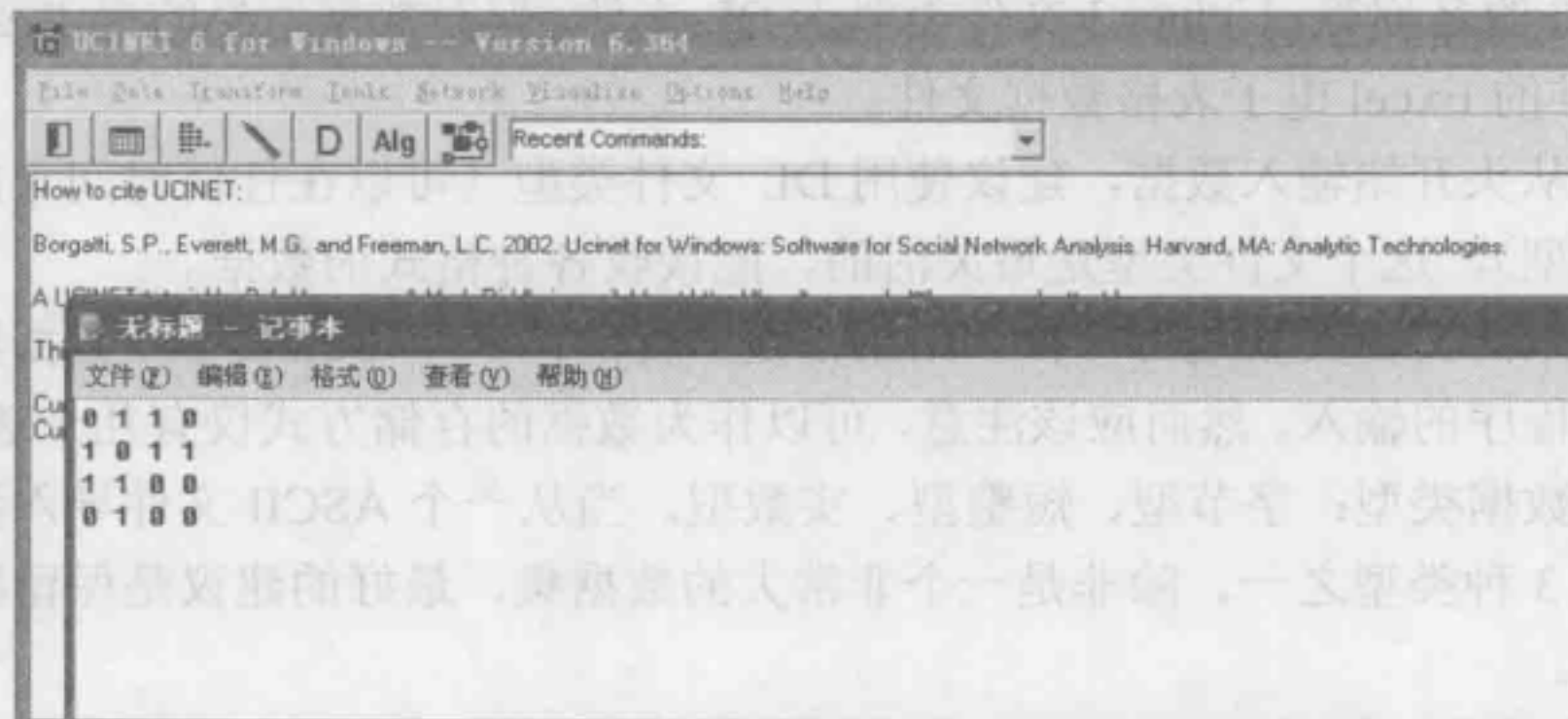


图 3-22 在文本编辑器中输入初始数据

输入完成后，可以直接保存为纯文本文件并重命名。然后在 Ucinet 中选择 Data | Import text file | Raw 命令，在出现的对话框中选择命名好的文件，就可以将文本文件中的数据转换成 Ucinet 数据。这种数据格式操作简单方便，但是建议一般不要使用。原因是，计算机程序不会对数据进行检查，如果数据的某一行缺少了一个值，程序本身不知道；数值没有标签，无法辨识不同的点。



### (2) 第二种数据格式

第二种数据格式是 Excel 文件数据。需要注意的是，Excel 数据文件虽然有上百万行，但数据的列数最多能支持 255 列。用户可以利用外部 Excel 软件输入数据，再使用 Import 命令将文件导入，如图 3-23 所示。也可以使用 Data | Spreadsheet 菜单或直接单击快捷菜单中的数据编辑器，打开后输入数据保存。

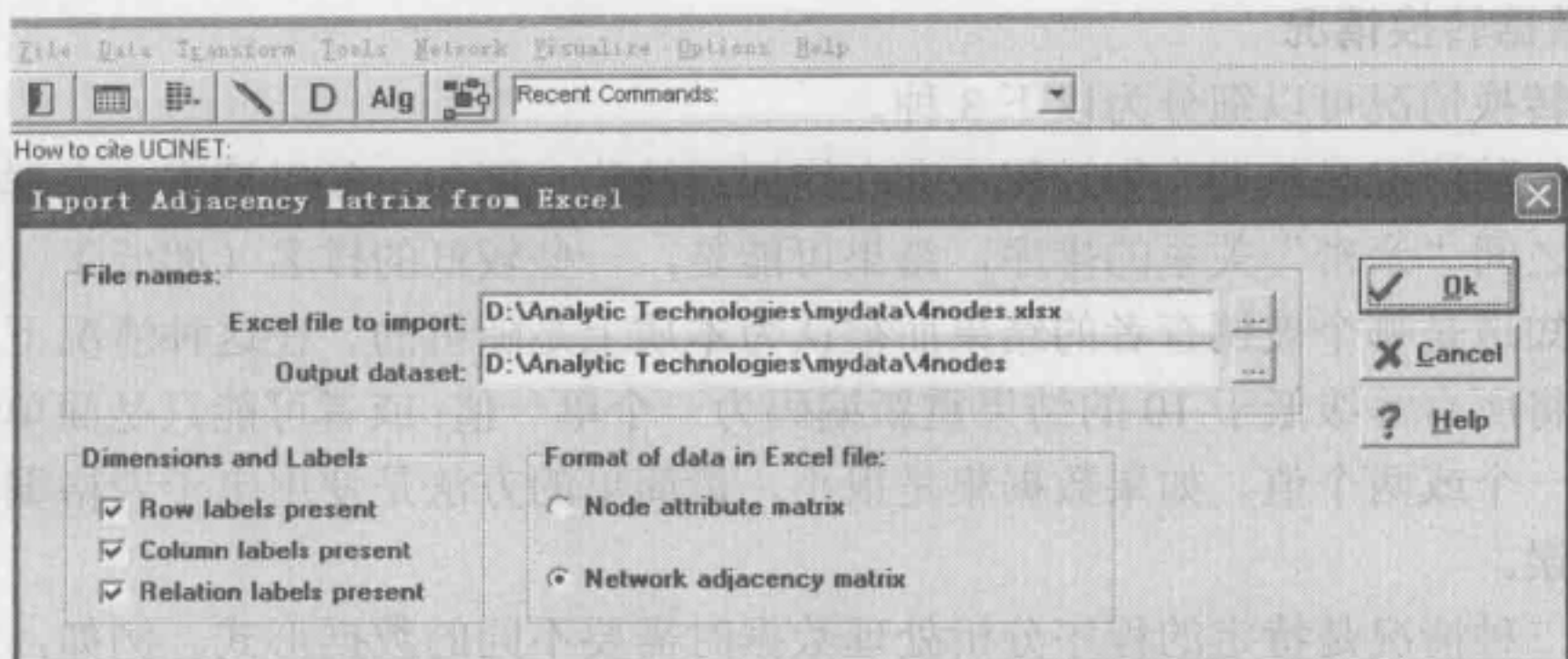


图 3-23 使用 Import 命令来导入文件

### (3) 第三种数据格式

第三种数据格式是数据语言文件 (DL File)，它包含一系列数字 (数据)、描述数据的很多关键词和语句等，这些都是关于数据的基本信息。数据文件正文的内容一定要用 DL 表明该文件是数据文件。在 Ucinet 文件中，常见的数据输入方法是在一个文本文件中输入数据，该过程可以利用任何一种文字处理程序 (如写字板或者 Word 等程序)。用 Ucinet 本身的 File | Text Editor 编辑器也可以输入数据文件，输入数据之后可以把数据保存为纯文本文件，然后利用 Data | Import text file | DL 命令就可以将文件转换为 Ucinet 本身的数据。在如图 3-24 所示的文件中，DL 表明是数据文件，n=4 表示矩阵是个四行四列的方阵；format=nodelist1 是矩阵的类型，即点列表格式，这种格式只能输入 1~0 矩阵数据，data 每行第一个数值 X 表示是矩阵的第 X 行，每行的后面的数值表示矩阵中哪个位置有数值。比如图中得到的矩阵为：

```
1 1 1 0
1 1 0 1
0 0 0 0
1 0 0 1
```

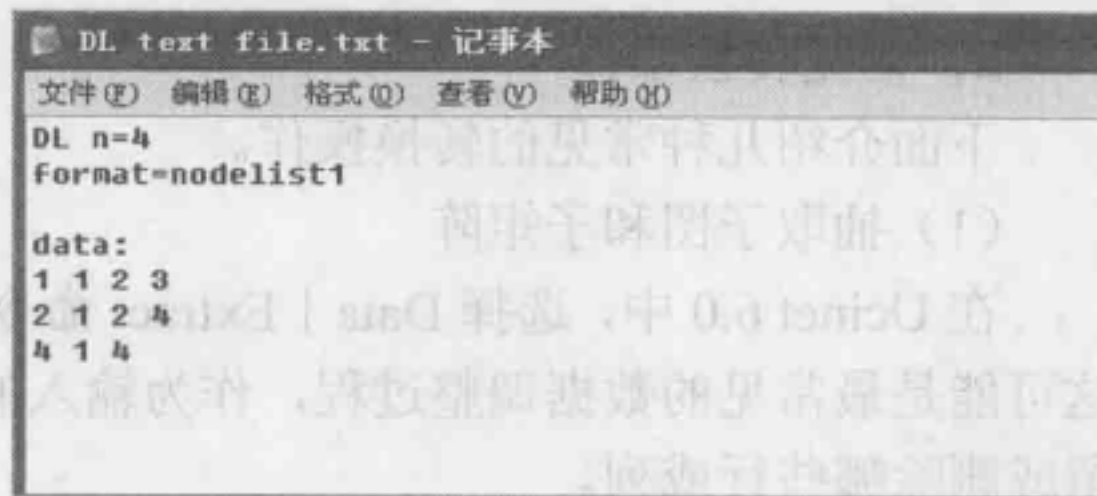


图 3-24 Ucinet 数据语言文件内容

DL 数据文件的格式还有全矩阵格式 (Full Matrix Format)、长方形矩阵 (Rectangular Matrices)、多个矩阵 (Multiple Matrices)、对角线缺失值矩阵 (Diagonal Absent)、左半矩阵或者右半矩阵 (Lower Half or Upper Half Matrices)、块矩阵格式 (Block Matrix)、关联列举格式 (Linked List Formats，该形式又包括点列表形式、等级列表格式、边列举格式等类型)。各个形式的矩阵都有表示矩阵类型的关键字和输入模式，需要了解这些矩阵格式的用户可以参考相关资料深入研究学习，这里不作详细介绍。



### 3.3.2 数据处理

有时想要对收集到的初始数据作进一步分析，就需要对初始数据做转换处理。数据转换的功能菜单上文已有介绍，此处对数据转换的情况进行简要说明，以帮助用户更好地理解数据转换功能，并能利用该功能得到自己所想要的的结果。

#### 1. 数据转换情况

数据转换情况可以细分为以下 3 种。

- 第一种情况是数据收集过程中存在错误或缺陷。例如，如果调查一个群体成员彼此之间“亲密”关系的排序，结果可能是，一些较低的排名（如低于 10），由于不知道是哪个被调查者的结果而被认为本质上是随机的。在这种情况下，可能决定将所有等级低于 10 的结果重新编码为一个单一值；或者可能只是简单地错误输入一个或两个值。如果数据集是很小，最简单的方法是使用电子表格编辑器改正错误。
- 第二种情况是特定的程序分析处理数据时需要不同的数据形式。例如，如果收集到多人之间的多值朋友关系数据，并且需要找出不同的派系，那么就需要将这些数据进行二分化，因为派系指令只能分析二值数据。
- 第三种数据转换的情况是需要创建新的变量，从而更便于分析。例如，假设收集到一群演员对一些特定问题的态度数据，同时也收集到不同演员之间的关系强度的数据。一个自然假设就是，检验态度上的相似与关系的强度是否有相关性。假设检验测试可以使用 QAP 程序，但首先需要获得行动者之间在态度上的相似性，这可以通过计算“行动者-态度”矩阵中隔行之间的相关性来确定。

Ucinet 6.0 可以执行各种各样的数据转换。然而需要注意的是，使用词语“转换”容易造成误导，因为没有一个是转换程序在实际上改变了现有的数据集。相反，是根据转换的目的生成了一个新的数据集（尽管这个新数据集可以覆盖现有的数据集），提交给程序进行分析操作的是刚生成的新的数据集，所以更像是一种深层次分析。

#### 2. 常见转换操作

下面介绍几种常见的转换操作。

##### (1) 抽取子图和子矩阵

在 Ucinet 6.0 中，选择 Data | Extract 命令，它会从一个输入矩阵中取出一个子矩阵。这可能是最常见的数据调整过程，作为输入的矩阵文件应该是多元的矩阵，需要指定保留或删除哪些行或列。

##### (2) 数据的合并

即将保存在不同文件中的同一类行动者之间的多种关系数据合并在一个文件中。比如对于某些程序来说（如 Positional Analysis），可以针对有关该数据的全部关系矩阵（如友谊关系、借款关系、亲属关系等）同时进行分析，这就需要将数据合并在一起。完成这项合并任务的指令是 Data | Join，合并的方式包括行合并、列合并、矩阵合并。

##### (3) 置换与排序

对矩阵的行和列重新排序，可以使同一个群组或同一类的成员互相邻近，以便更



方便地了解矩阵关系的整体结构。Ucinet 6.0 给出了两种基本排序类型，即置换 (Permutations) 和排序 (Sorts)。置换是用户直接指定结点的排列顺序。例如，可以指定程序重新排列结点的要求为“1, 6, 2to5, 10, 9, 8, 7”，这句指令的意思就是想让结点 1 排列为第一位，结点 6 排第二，以此类推。

(4) 对称化处理

当收集的关系数据本身具有对称性，如“与谁共进午餐”，这种关系自身带有对称性，或者一些非对称的关系拓展成为对称关系，如将“借钱给某人”关系拓展成“金钱关系”，在这种情况下，需要将关系矩阵对称化处理 (Symmetrizing)，即对矩阵的全部  $i$  和  $j$  来说，令  $X(i, j)$  等于  $X(j, i)$ 。

(5) 标准化处理

数据转换程序的标准化处理 (Normalizing)，也是常用的一种指令。它可以对整个矩阵数据或者某些行和列进行标准化处理，如要求一个矩阵的各行都有同样的均值或者标准差等。特别是在矩阵的数据收集是由不同的回答者得出的结果，这种处理的要求更迫切。比如，如果受访者被要求评估从家里到家庭网络中的所有其他成员之间的物理距离，也许一些受访者用多少码来回答，而另一些受访者可能以多少米来回答。在这种情况下，就有必要将数据进行标准化，使得数据有可比性。

数据处理还有多种转换矩阵格式的方法（如转置和改型、重新编码、改变模态等），通过数据处理后不仅可以得到分析程序要求的数据格式，还可以对数据进行深层挖掘，深入分析。

### 3.4 Ucinet 应用举例

本节以主题“手机图书馆”为例，利用 Ucinet 软件的主要功能来实现相关图谱创建，包括数据处理与图谱创建两部分。图谱创建部分针对软件主要功能介绍实现的过程，并对形成的图谱涵义解释说明。

本案例将以 CNKI 数据库为数据源，结合文献计量分析理论和社会网络分析方法，利用 Ucinet 软件对新兴的手机图书馆领域的几个方面进行举例分析，通过对高产作者和高频关键词特征等方面分析，对手机图书馆领域的发展状况进行阐述，并将作者共现矩阵、高频关键词共现矩阵导入 Ucinet 中进行可视化分析。

#### 3.4.1 数据处理

本案例在 2013 年 4 月 24 日搜索 CNKI 数据库总库，以“手机图书馆”、“移动图书馆”为关键词，分别在主题、关键词、题名字段进行检索，得到的检索结果为 313 条。为了浏览数据方便，选择每屏显示 50 条数据，逐条查看筛选并标记相关文献并选择存盘，选择 Endnote 格式，将标记好的文献题录保存为后缀为.txt 的文本文件。数据清理后共有 309 条相关文献的题录信息，共 7 个.txt 文件，并将此 7 个文本文件存储于新建文件夹 data 中。

由于需要对作者共现和高频关键词共现进行可视化分析，所以需要原始的数据进行处理，生成 Ucinet 软件可以分析的共现矩阵文件。案例选用了浙江大学信息资源管理





色节点是度数为 0 的著者，即没有出现共现的作者群，而右侧有连线的著者群是相互之间有共现的作者群，连线的粗细代表两个作者之间联系的强度。

从图中可以看出，44 位高产作者形成的合著网络图谱整体联系稀疏，仅有少数几位作家形成小团体。第一个小团体是成都理工大学及成都市图书馆为合作机构下的程孝良、李勇、钟刚毅、叶艳鸣 4 位作者合著，合著频率是 4 次，在该合著网络中频率达到最高。该团体主要针对手机图书馆的传播影响和设计理念进行了多次探讨。第二个小团体是南京农业大学的茆意宏、黄水清、吴政 3 位作者合著的群体，其中代表茆意宏的节点最大，表明在该团体中的频次最高，在三者中占有相对重要的地位，主要研究手机图书馆的兴起、发展、系统建设、服务需求等。第三个小团体是重庆邮电大学的田增山、李嫒源、贺利娜 3 位作者形成的群体，作者之间共现频次是 2 次，主要对手机图书馆系统的构建做了论述。网络中大多数节点之间没有关系，相互之间的关联程度非常小，表明正在发展的手机图书馆领域还没有一支成熟的研究队伍。

## 2. 合著网络密度分析

网络密度是社会网络分析中常用的一种测度，网络密度反映的是节点间联系的紧密程度。在 Ucinet 中，选择 Network | Cohesion | Density 命令，可计算作者合著网络的密度，得到如图 3-26 所示的界面。

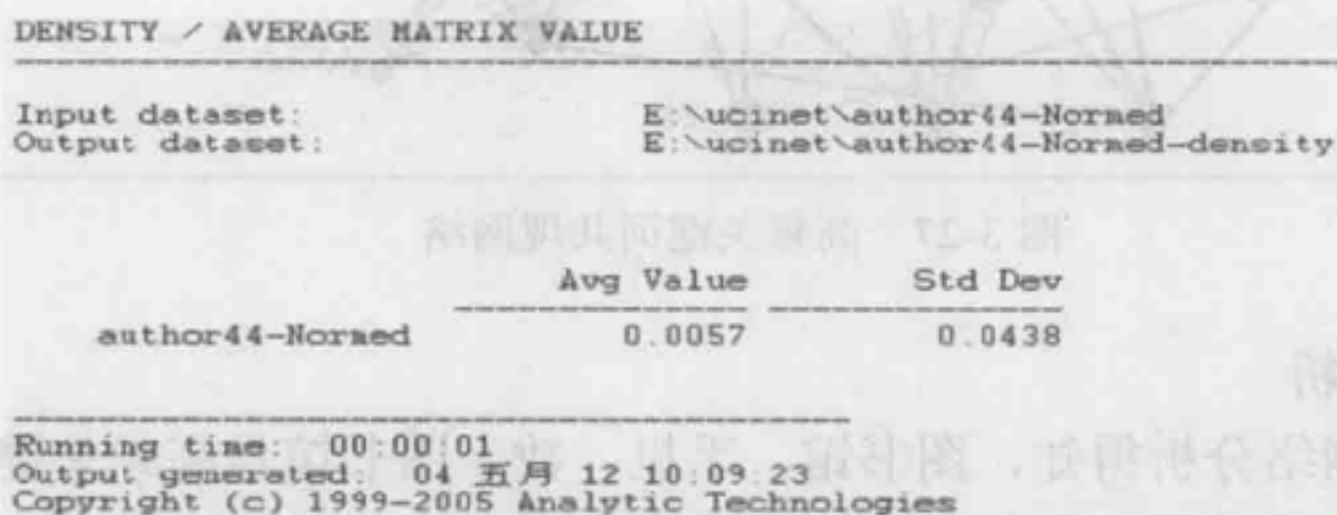


图 3-26 合著网络密度结果

从图 3-26 中可以看出，作者合著网络密度为 0.0057，表明节点之间的紧密程度较低，各个节点之间的连线较少。通过作者字段数据的整理，发现大多数著者为独立作者，仅有少数作者之间合著发文。因为手机图书馆领域正处于发展时期，多数作者对该领域的研究还处于起始阶段，各个机构下的作者几乎没有相互合作，所以相互之间的合著几率非常小。

## 3. 高频关键词网络分析

同合著者网络的处理办法一致，将高频关键词的共词矩阵导入到 Ucinet 中，生成高频关键词共现网络，利用关键词之间关系的紧密程度分析出手机图书馆领域的主要研究内容。

打开 Netdraw 软件，导入关键词矩阵，选择 Analysis | Centrality Measures | Betweenness 命令，得到基于节点在网络中作用大小的图谱。

图 3-27 中的圆点代表高频关键词节点，节点越大，代表节点在网络中的作用越大。节点之间的关系用实线连接，实线越粗，代表相互之间的关系越强。从图中可以看出，高频关键词网络以手机图书馆、移动图书馆为核心，这两个节点与其他节点的联系最紧密，是该领域的核心关键词；其次，比较重要的节点还有图书馆、手机、数字图书馆，这 3 个节点在网络中也起到了比较重要的作用，与其他关键词有较强的联系，能够反映出该

领域的研究内容。从整个网络来看，核心关键词以外的结点联系较为松散，说明手机图书馆领域还处于发展时期，还没有形成集中的研究主题。为了更深入地探析目前手机图书馆研究的主要内容，下面进一步分析除核心关键词以外的 3 个主要关键词的网络特征。

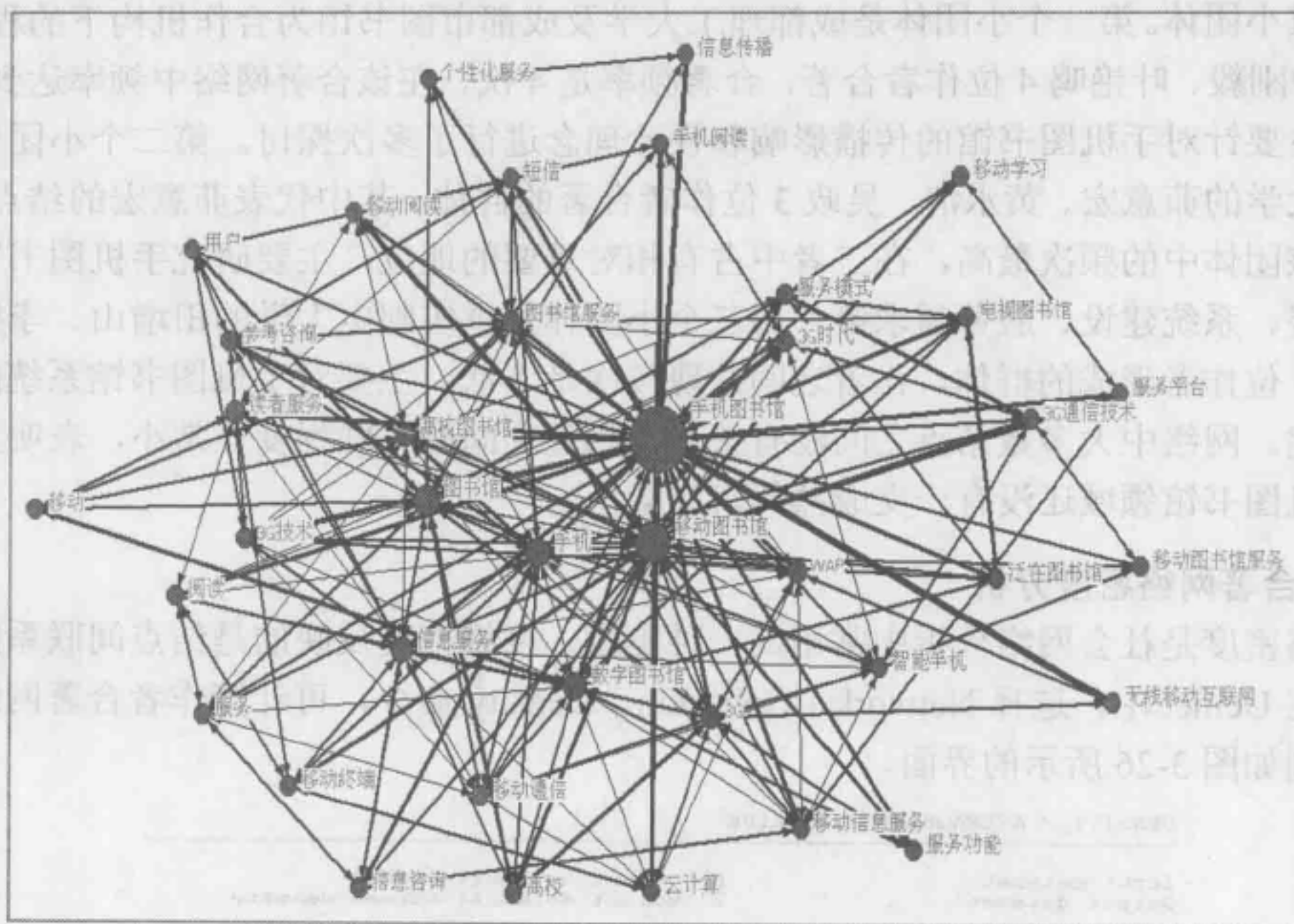


图 3-27 高频关键词共现网络

#### 4. 个体网络分析

由高频关键词网络分析得知，图书馆、手机、数字图书馆是高频关键词网络中的主要关键词，通过个体关键词的网络特征可以清晰发现结点间的更多关系。选择 Netdraw Analysis | Structural Holes | Ego Network Model 命令，选择各个主要关键词的个体网络。选取关键词“图书馆”的个体网络为例，如图 3-28 所示，其他个体网络图谱省略。

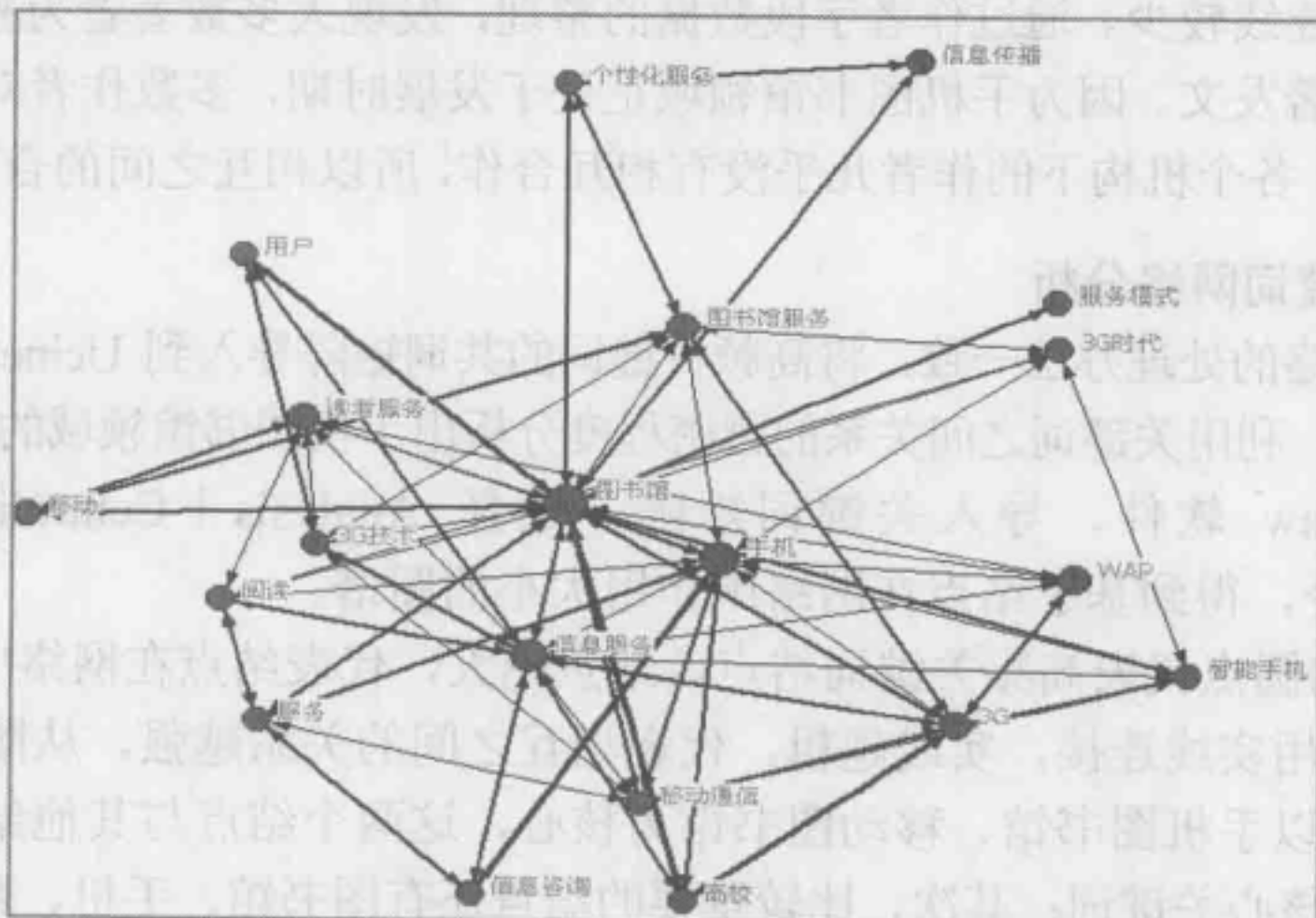


图 3-28 关键词为图书馆的个体网络



从图 3-28 中不难看出,与关键词“图书馆”相连接最多的关键词有:高校、用户、手机、个性化服务、移动、服务、信息服务、服务模式、读者服务、图书馆服务等。表明手机图书馆领域主要的研究机构是高校图书馆,研究的宗旨是围绕使用手机提供图书馆的个性化服务,提出新环境下的图书馆提供的服务模式,更好地满足读者的需求。

选择关键词“手机”的个体网络,与其连接最多的结点是:信息咨询、参考咨询、服务、移动图书馆服务、移动通信、3G 通信技术等。可见手机图书馆领域的一个重要研究方向是使用手机提供图书馆信息咨询服务的途径,不仅注重图书馆的服务内容,也要考虑如何利用移动通信技术提供手机信息参考咨询服务。

选择关键词“数字图书馆”的个体网络,与其连接最多的结点有:移动、信息服务、移动通信、服务功能、服务模式、泛在图书馆、高校、3G 技术、云计算等。表明数字图书馆在手机图书馆的发展中起到非常重要的作用,需要就图书馆如何实现移动服务提供支持。

从文献高频关键词统计结果看出,手机图书馆的研究内容较为宽泛,与该领域相关的图书馆、信息服务、3G 技术、手机、短信等都有提出,但还尚未形成明显的研究主题,研究方向还处在探索阶段。



## 本章习题

1. Ucinet 软件的功能包括哪几部分?
2. 在实际操作中, Ucinet 的运行速度与哪些因素有关?
3. Ucinet 可以读取哪些类型的文件?
4. Ucinet 6.0 程序运行的两种方式是什么?
5. Ucinet 软件的主要功能菜单包括哪些?
6. Ucinet 输入与输出的文件格式有哪些类型?
7. 数据分块是指什么? 请举例说明。
8. 网络密度分析有几个指标?
9. 网络中心度分析中, 度数表示结点的什么问题?
10. 在文献计量研究中, 使用 Ucinet 及 Netdraw 软件绘制知识图谱有哪些优势?
11. Ucinet 可选择的可视化软件有哪些? 各自有什么特点?
12. 列举 3 个 Netdraw 软件中调节知识图谱的指令。
13. Ucinet 支持的数据格式有哪几种?
14. Ucinet 中数据处理的方式有几种?
15. 什么情况下需要进行 Ucinet 数据转换?
16. Ucinet 数据标准化有什么意义?
17. Ucinet 数据处理的目的是什么?
18. 使用 Ucinet 进行数据分析需要注意哪些问题?
19. 使用 Ucinet 分析数据前, 使用其他软件处理数据是否有必要?
20. 利用 Ucinet 知识图谱功能, 可以对某一知识领域的哪些方面进行分析?

