这本书被称为复杂网络入门读物中最全面的一个，由M.E.J.Newman编著，（Newman的主页ww-personal.umich.edu/~mejn/）

介绍

网络的科学研究，比如计算机网络、生物网络和社会网络，是一个跨学科的领域，包含了从数学、物理学、生物学、计算机科学、社会科学、和许多其他领域的思想……

书分为5章。第一章描述了网络的基本类型，这些类型用于研究现今的科学和用于判断他们的结构的实验技术。Part I describes the basic types of networks studied by present-day science and the empirical techniques used to determine their structure.第二章介绍了基本的数学工具，以及方法和量化网络结构的统计学。Part II introduces the fundamental mathematical tools used in the study of networks as well as measures and statistics for quantifying network structure.第三章描述了有效分析网络数据的计算机算法。Part III describes computer algorithms for the efficient analysis of network data,第四章描述了网络结构的数学模型，用于预测网络化的系统的行为并且理解他们的形成和演化while Part IV describes mathematical models of network structure that can help us predict the behavior of networked systems and understand their formation and growth. Finally, Part V describes theories of processes taking place on networks, such as epidemics on social networks or search processes on computer networks.

……一通感谢的话……

目录

第一章——介绍

我们为什么对网络感兴趣

网络的一些例子

网络的属性

这本书的大纲

第一部分——网络的实证研究

第二章——技术网络

2.1 互联网

2.2 电话网

2.3 电网

2.4 交通网

2.5 DELIVERY AND DISTRIBUTION NETWORKS

第三章——社会网络

3.1 社会网络的实证研究

3.2 问卷调查

3.3 直观结论

3.4 档案和第三方记录的数据

3.5 联盟网络

3.6 小世界实验

3.7 雪球采样，接触追踪CONTACT TRACING，随机游走

第四章——信息网络

4.1 WWW

4.2 引文网络CITATION NETWORKS

4.3 其它信息网络

第五章——生物网络

5.1 生化网络

5.2 神经网络

5.3 生态网络

第二部分——网络理论基础

第六章——网络中的数学

6.1 网络及其他们的表述

6.2 邻接矩阵

6.3 权重网络

6,4 有向网络

6.5 超图（网）

6.6 二分图（网）

6.7 树

6.8 平面网络

6.9 度

6.10 路径

6.11 组件

6.12 INDEPENDENT PATHS独立通路, CONNECTIVITY连通性, AND CUT SETS割集

6.13 拉普拉斯算子

6.14 随机游走

问题

第七章——MEASURES AND METRICS指标

7.1 中心度

7.2 EIGENVECTOR CENTRALITY中心特征向量

7.3 KATZ CENTRALITY

7.4 PAGERANK

7.5 HUBS AND AUTHORITIES

7.6 接近中心性

7.7 BETWEENNESS CENTRALITY介数中心性

7.8 顶点集

7.9 传递性

7.10 互反性RECIPROCITY

7.11 SIGNED EDGES AND STRUCTURAL BALANCE

7.12 SIMILARITY相似性

7.13 同质性和相称混合性

问题

第八章——大规模网络

8.1 组成

8.2 最短路径和小世界效应

8.3 度分布

8.4 幂律和无标度网络

8.5 其它中心度量的分布DISTRIBUTIONS OF OTHER CENTRALITY MEASURES

8.6聚类系数

8.7相称混合性

问题

第三部分——计算机算法

第九章——算法的基本概念

9.1 运行时间和计算复杂度

9.2 存储网络数据

9.3 邻接矩阵

9.4 邻接表

9.5 树

9.6 其它网络表述

9.7 堆

问题

第十章——基本网络算法

10.1 度和度分布算法

10.2 聚类系数

10.3 最短路径和广度优先搜索

10.4 网络不同边长度的最短路径

10.5最小截和最大流

问题

第十一章——矩阵算法和图分割

11.1 LEADING EIGENVECTORS主要特征向量 AND EIGENVECTOR CENTRALITY

11.2 网络聚类

11.3 图的划分

11.4 THE KERNIGHAN-LIN ALGORITHM

11.5 谱划分

11.6社区发现

11.7 SIMPLE MODULARITY MAXIMIZATION

11.8 SPECTRAL MODULARITY MAXIMIZATION

11.9 划分到两个以上的组（多重划分？）

11.10 OTHER MODULARITY MAXIMIZATION METHODS

11.11 其它社区发现算法

问题

## 第一章 介绍

为什么我们对网络感兴趣？

许多都是网络

网络中成员之间连接的模式很重要

这个模式可以用网络表述，网络的结构和模式对系统的行为影响很大

网络是系统的简化

最普通的网络变体在第六章讨论

科学家使用数学的，计算机的，统计学的工具去建模去理解网络。

网络的应用范围很广

一些网络的例子

描述Internet……

Internet在2.1节详述

Internet的一些研究点

Internet的结构在新的通信标准的发展中有重要地位，对它更好的理解可以提高表现

图1.1：Internet网络结构C类子网的Internet，在2.1节讨论

Web和Internet这两个词经常被人们认为相互可替代，但从技术层面说它们大有不同。Internet是一个物理的网络，用真实的线或无线电在电脑之间连接。而Web，是存储在网页上的信息的网络。WWW网络的顶点是网页，而边是超链接，这里是没有物理结构的。

WWW在4.1节讨论

WWW网络的研究价值

网络搜索的机制在19.1节讨论

WWW为有向网。

社会网络在第三章做深度探讨

描述社会网络……

一些社会网络的例子。

图1.2：一个俱乐部中会员的朋友网

神经网络在5.2和5.3节探讨

比较重要的第三个领域是生物。比如神经网络，图1.3是一个“食物网络”，刻画了一个生态网络，顶点是一个生态系统中的物种，边表示捕食-被捕食关系，这也是生物网络的一个分支。

图1.3：食物链网络

生化网络在5.1节探讨

另一种生物网络是生化网络，比如新陈代谢网络，蛋白质相互作用网络，基因调控网络。

以上仅为几个本文关注的几个类型的网络的例子，还有一些其他的比如，电话网，道路、铁路、空中交通网，电力网，引文网络，recommender networks，点对点网，电子邮件网，合作网，疾病传播网，河流网，word networks。

网络的性质PROPERTIES OF NETWORKS

对于一个网络，如果我们获得了它的数据，我们可以对这些数据做什么？关于这个网络所表述的系统的形成和功能，这些数据能告诉我们什么？网络化的系统有哪些性质可以供我们测量或者模拟，并且这些问题与我们关系的问题有哪些关联？

观察网络结构的第一步往往是画个图，但是对于大规模网络就行不通了，网络科学自有一套方法。

一类非常有用的方法是对中心centrality的测量。比如：度。

第七章进一步讨论中心测量。

有向网中有入度和出度两种

可以观察到一个有趣的现象，许多网络中都有一些小但有相当数量的“枢纽hubs”——度数高的顶点。许多实证和理论研究表明“枢纽hubs”具有相当不成比例的影响，playing a central role particularly in network transport phenomena and resilience, despite being few in number.

Hubs将进一步在8.3节讨论

小世界效应……小世界效应是另一个不断被提到及应用的概念

小世界效应例子，流言一眨眼间便可传播整个网络

小世界Internet的例子

第三个不断被提到及应用的例子是簇或者说社区。……这些是目前高度活跃的主题，有希望在未来应用到实践中

OUTLINE OF THIS BOOK

This book is divided into five parts. In the first part, consisting of Chapters 2 to 5, we introduce the various types of network encountered in the real world, including technological, social, and biological networks, and the empirical techniques used to discover their structure. Although it is not the purpose of this book to describe any one particular network in great detail, the study of networks is nonetheless firmly founded on empirical observations and a good understanding of what data are available and how they are obtained is immensely极大地 helpful in understanding the science of networks as it is practiced today.

The second part of the book, Chapters 6 to 8, introduces the fundamental theoretical ideas on which our current understanding of networks is based. Chapter 6 describes the basic mathematics used to capture network ideas, Chapter 7 describes the measures and metrics度量 we use to quantify network structure, and Chapter 8 describes some of the intriguing有趣的 patterns and principles that emerge when we apply our mathematics and our metrics to real-world network data.

In the third part of the book, Chapters 9 to 11, we discuss computer algorithms for analyzing and understanding network data. Measurements of network properties, such as those described in Chapter 7, are typically only possible with the help of fast computers and much effort has been devoted over the years to the development of efficient algorithms for analyzing network data. This part of the book describes in detail some of the most important of these algorithms. A knowledge of this material will be of use to anyone who wants to work with network data.

In the fourth part of the book, Chapters 12 to 15, we look at mathematical models of networks. The material in these chapters forms a central part of the canon标准 of the field and has been the subject of a vast amount of published scientific research. We study both traditional models, such as random graphs and their extensions, and newer models, such as models of growing networks and the “small-world model.”

Finally, in the fifth and last part of the book, Chapters 16 to 19, we look at processes taking place on networks, including failure processes and resilience回弹, network epidemiology流行病学, dynamical动力学 systems, and network search processes. The theory of these processes is less well developed than other aspects of the theory of networks and there is much work still to be done. The last chapters of the book probably raise at least as many questions as they answer, but this, surely, is a good thing. With luck readers will feel inspired to answer some of those questions themselves and the author looks forward to the new and exciting results they generate when they do.

第一部分 网络的实证研究

## 第四章 信息网络

### 4.2 引文网络

一个比较名不见经传但是比较老的信息网络是学术论文间的引文网络。许多论文参考了一个或者多个之前的论文，多见文章底部的文献目录，由此可以构造出一个网络。有许多原因为什么一个文章会引用另一个——为了之处对读者有用的信息、为了赞扬之前的工作、为了指出对目前工作的影响、或者不同意文章的内容。概括地说，不管怎么样，如果一篇文章引用其他文章通常说明这个文章的内容在某种程度上和自己相关，因此引文网络是一种和主题关联的网络。

引文网络的量化研究可追述到1960s。最早的可能是1965年Price [274]（同时也是我们已知的最早的发现幂律度分布power-law degree distribution的研究，我们将在8.4节详细讨论）。这些研究有时被归为“图书科学”，但是现在经常被称为“信息科学”。专门处理出版和引用的统计研究的信息科学的分支被称为bibliometrics文献计量学。

收集引用数据。Price收集出了一个SCI……

见4.1节网页爬虫的讨论

最近，电脑自动的引文索引变得很普遍。比如，网站Citeseer，由Pennsylvania State University维护，提供论文引用索引信息。这是一个某种的无目的的操作hit-ormiss operation，因为许多文章不在网上并且不是免费的，并且文章中的引用有不同的格式并且会有错误，而且一篇文章可能在网上存在多个地方，在期刊和书记上也是，而且可能有多个版本。尽管如此。Citeseer还是很好用的。还有其他的automatic citation indexing projects包括Citebase 它indexes physics papers还有Google Scholar。

正如网页爬虫…………………………没看懂这段想讲啥…………………………

引文网络在很多方面和WWW相似。网络的结点用文字和图片的方式保存信息，和一个网页是一样一样的，连接的方式和网页上超链接也是相似的。具有很多引用的文章通常具有更大的影响以及更广阔的阅读量，和网页正相似。读者可以通过引用找到参考文献和网页也是一样的……

Acyclic networks非循环网络在6.4.2节讨论

引文网络和web有一个重要的不同点：引文网络是非循环的，而web不是。……在6.4.2节我有详细说明……

见图6.3一个简单的非循环网络插图

引文网络Citation networks有一些令人惊奇的统计。SCI中大约47%的文章从未被引用过。剩余的，9%被引用过一次，6%两次，之后的就下降的非常厉害。只有21%的文章被引用过10次或以上，只有1%的文章被引用100次或以上。这指出一个重要结论：网络的幂律度power-law degree distribution分布，在8.4节有详细说明

在SCI上引用最高的文章是Lowry et al的一篇分子生物学的方法论的文章。

引文网络是最简单的但并不是引用模式citation patterns里最具有代表性的网络。另一个被广泛研究的表现形式是cocitation network。当两篇文章同时被另一篇文章引用时他们被称为cocited，cocitation表示了文章的主题相似度。

weighted cocitation networks这些在后面都有介绍。

另一个相关的概念，但也并不常用，是bibliographic coupling。两篇文章都引用了另一篇文章时被称为bibliographically coupled。……在后面也有……

Cocitation and bibliographic coupling are discussed in more detail in Section 6.4.1.

### 4.2.1 PATENT AND LEGAL CITATIONS专利和法律引用

我们对引文网络的讨论目前还集中在学术论文之间的引用，但是还有其他类型的引用。有两种比较重要的是专利间的引用和法律意见书间的引用。

专利是…………解释专利…………。专利的重要问题：政府通过审查程序判断这个发明是原创的还是之前就有的。对于一个专利的申请，发明人以充足的理由必须描述他的发明的价值值得专利保护。这个例子典型的包括了处理发明和前人发明的关系，为了这么做发明人通常会引用一个或多个之前的专利。引用可能会高亮两个技术直接的相关性，比如一个发明依据于他对另一个发明的操作，但是通常专利的引用是“defensive”，意思是发明人引用之前技术的专利，然后提出一个论点为什么新的技术与老的技术足够不同以突出他自己的专利。政府在检测专利申请的过程中时，会例行公事地考虑它与之前专利的相似性，那么defensive citations是一个发明人可以预先避开可能的冲突的方法。代表性地，政府专利检验员和发明人在专利申请被最终接受或拒绝之前，会有数轮交流。在这个过程中，额外的引用经常会被发明人或者检测员加入到申请中，去用文件记录他们交流中讨论的点。

如果一个专利最终被认可，之后就是发布，包括引用等等，这样大众就会知道哪些技术是被保护的。这些发布的专利提供了我们可以用来构建网络的引用数据源。在这样的网络中顶点是专利，由专利号标识，之后连接……。专利网络大多是非循环的……但是有个别例子，发明人同时地申请一些相互独立的技术的专利。专利网络的结构反射出the organization of human technology。专利网络的研究远不及学术引用网络，但是由于高质量数据集的出现最近几年增长很快，

第二部分 网络理论基础

## 第六章 网络中的数学

一个对网络研究中使用的数学工具的介绍，这些工具在许多随后的发展中会很重要。

在接下来的三章我们会介绍fundamental quantitative foundations。在这章中我们介绍用于描述和分析网络的基础理论工具，许多都来自图论。

第七章是量化网络结构的方法和度量metrics；第八章，对于一些出现在真实世界的网的标志性的模式，我们将数学和metrics使用到它们的分析上。

### 6.1 网络及其他们的表述

不同的领域顶点和边的表述不一样，比如nodes and links in computer science, sites and bonds in physics, and actors43 and ties in sociology

表6.1：不同的领域的不同表述

本书一般n表示顶点的数量，m表示边的数量

本书大部分网络中两顶点间只有一条边，非多边，并且不自连（self-edges or self-loops）。

无自连和多边的网络成为简单网络或简单图。否则称作……

### 6.2 邻接矩阵

一个网络……

用边集/表（edge list）表示，不过不推荐

图6.1：两个网络，a是简单图，b包含了多边和自连接

一个更好的表述方法是邻接矩阵

Aij=1（顶点i和j之间有边）=0（无边）（6.1）

一个邻接矩阵的例子（6.2）

两点注意：（1）无自连接的网络，主对角线为0，（2）矩阵为对称矩阵

多边和自连接亦可用邻接矩阵表述。多边用边数填充矩阵中位置

自连接，注意的是，主对角线上数值是2，为什么是2……

另一种为什么是2的理解方式是，对于非自连接的一条边，在矩阵中有两个地方是1，所以自连接以2表示

（6.3）图6.1中b的邻接矩阵

还有一种自连接的多边（自己连自己多次），连x次，则主对角线位置2x。

### 6.3 权重网络

权重网络……

（6.4）一个权重网络的邻接矩阵

如果一个权重网络都是实数，那么有时候可以适当转化为多边网络有利于分析，反之亦然

网络上的权重亦可是负值，比如在社会网络中表示憎恨。这会进一步在7.11节讲述，我们会考虑到结构平衡structural balance的概念

边可以有权重，顶点也可以有，权重还可以是向量和离散变量比如颜色。

### 6.4 有向网络

有向网络……

图6.2 一个有向网络

前面已举例了很多有向网络，超链接组成的有向网络、食物网、引用网

（6.5）Aij=1（有一条边从j到i）=0（无边）（6.1）

注意：Aij=1表示的边的方向是从第二个顶点到第一个顶点。这有点反人类，但在数学上却是很方便的。

（6.6）图6.2的邻接矩阵

注意到有向图的邻接矩阵不一定对称

无向图也可看作有向图，其每条边变成了两条有方向的边

同样的，有向图也可以多边和自连接

#### 6.4.1 COCITATION AND BIBLIOGRAPHIC COUPLING

出于分析的需要有时候需要将有向图转换为无向图，因为无向图有许多有用的分析工具而有向图没有。

一个简单的方法是直接无视有向图边的方向，但不可避免地抛弃了一些网络结构的潜在信息。一个复杂的方法是使用“cocitation” or “bibliographic coupling”，两个不同但是相关的思想，他们的名字来自于它们被广泛使用的引文网络分析领域。

我们在4.2节简单地讨论过cocitation

有向网络中，两个顶点i和j的cocitation值是同时指向i和j的顶点数。比如在引文网络中，2篇文章的cocitation是其它同时引用这两篇文章的文献的数目。易知AikAjk = 1，当顶点i和j同时被k引用时，否则为0。将所有k加起来，得到cocitation Cij of i and j is（6.7）

这样（6.8）cocitation矩阵

容易证明C是对称矩阵

这样我们就可以建立一个cocitation网络，当Cij > 0时，这个网络顶点i和j之间有边，表示i和j在原来的有向网中被共同引用。这样cocitation网络就是一个权重网络，权值为正整数Cij，并且是一个无向网。

cocitation网络很有意义，比如，两个文献间cocitation很大（权值很大、连接很强）对论文是一个很好的指示，去处理相关的主题——如果两篇文章经常在同一个参考目录中被引用，他们可能在某些地方有相似之处。并且越频繁，越相关。

注意的是：cocitation矩阵的对角元素如6.9所示

（6.9）

即网络中总共引用i的次数。但在构建cocitation网络的时候，我们将对角元素设为0.

顶点i和j引用了三篇同样的文献，所以bibliographic coupling为3

bibliographic coupling和cocitation相似。bibliographic coupling是两个顶点指向相同顶点的数目，如k被i和j同时引用，则记AkiAkj = 1，否则为0.同理

(6.10) 

(6.11) 

bibliographic coupling矩阵同样是对称矩阵，并且非对角元素对应于无向网络的权重，

(6.12) 

Bii对角线元素就是文章i引用的文献数量

Bibliographic coupling也和cocitation一样有用。比如，如果两篇论文引用了许多相同的文献，这表示着他们在同一个课题……

虽然cocitation and bibliographic coupling数学上是相似的方法，但是在实际中却会产生不同的结果。尤其是，他们受出度、入度数量影响很强。对于两个顶点要让他们有强cocitation，他们首先必须有很多入度。在一个引文网络，比如，两篇文章要有强cocitation则必须被广泛地引用，这样一来，有强cocitation的文章仅限于有影响力的文章，比如综述文章，书，以及其他一些类似的高引文章。相反的，两篇文章要想有强bibliographic coupling则他们必须共同引用很多其他文章。在实践中，参考文献的规模远小于收到的引用文章的数目，那么bibliographic coupling是一个分析两篇文章相似度的更一般的方法。比如SCI，就是使用bibliographic coupling到他们的“Related Records” 特征，可以让用户找到相似的文章。而cocitation在这种情况下就不推荐，因为它在文章被引用很少的情况下表现不好。

Bibliographic coupling还有一个优点就是文章一发表就可以计算。Cocitation得等着被引用。另外，随着文章逐渐地被引用，Cocitation会发生变化，Bibliographic coupling不会，但是这点既可以说是优点也可以说是缺点，因为这样Cocitation会发现些有趣的性质，而如同Bibliographic coupling一样不变的方法则不会。

这两种方法除了用于衡量顶点的相似度外，还被用于有向网的搜索算法，被称为HITS算法，将在7.5节介绍。

#### 6.4.2 非循环的有向网络

网络中一个循环如图所示。类似WWW的网络有很多这样的循环。没有循环的网络就叫做acyclic非循环 networks。注意：一个自连接的顶点也算一个循环，所以非循环网络没有自连接。

一个经典的非循环网络的例子就是引文网络，如同4.2节所讨论的那样。在写文章的时候，只可能引用已经存在的文献，这表示着在引文网络中所有有向边指向以前的时间。这样我们如图6.3那样画图，将顶点按照时间先后从小到大给标签赋值，并且标签越小，越在图的底下，这样图中所有箭头都是朝下的。很容易理解这是没有循环的。……证明略……



图6.3：一个非循环有向图。

很容易观察到如果一个网络是非循环的，那么可以画出一个图，所有边都是指向下的。这个证明是很有用的，因为它同时提供给我们一个方法去判断一个网络是否是非循环的。

假设我们有一个n个结点的非循环的有向网络，那么在这个网络中必然有一个节点只有入度没有出度。……证明略……。

在实践中，不必去画出一条路径去找到无出度结点，我们只要知道存在即可。

我们现在画出无出度结点及附着着它的边，然后顺着边，找到去掉这个边变成无出度结点的结点，画在上一批的上面，依次往复。

当我们画完所有的结点，然后再将边加上，就形成了所有边朝下的网络。此网络不唯一

这个过程被很多算法所用。

这个过程也可以用于判断是否有循环，如果一个网络有循环，那么在此过程中一定有不能删除的结点。所以一个简单的判断网络是否循环的算法如下：

1. Find a vertex with no outgoing edges.

2. If no such vertex exists, the network is cyclic. Otherwise, if such a vertex does exist, remove it and all its ingoing edges from the network.

3. If all vertices have been removed, the network is acyclic. Otherwise go back to step 1.

非循环有向网络的邻接矩阵有很有趣的属性。比如上图的网络，我们画出它的邻接矩阵（6.13）它的下三角形全是0，……原因不说，很容易……

主对角线也是0，不题。值得说一下，对角线为0 的上三角矩阵称为strictly triangular

当然网络的邻接矩阵并不唯一，但是我们可以说必定存在邻接矩阵是严格上三角。

邻接矩阵还有一个性质就是，它的特征值都是0当且仅当网络是非循环的。为了证明这个……

以下是证明……

略

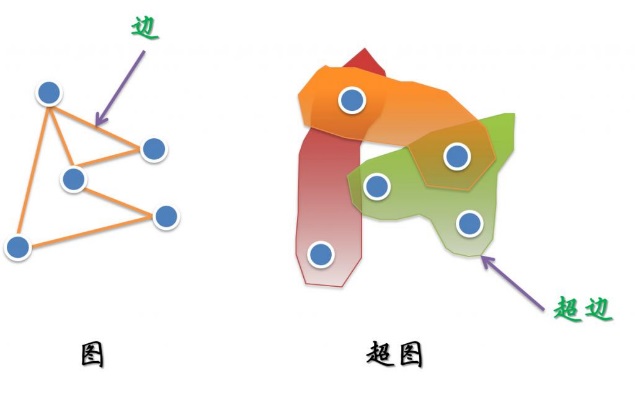
特征值都是0的矩阵叫做幂零矩阵。因此我们可以说一个网络是非循环的当且仅当它存在一个幂零邻接矩阵。

### 6.5 超图

！非翻译\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

超图（Hypergraph）是什么

简单的来说，对于我们熟悉的图而言，它的一个边（edge）只能和两个顶点连接；而对于超图来讲，人们定义它的边（这里叫超边，hyperedge）可以和任意个数的顶点连接。一个图和超图的示意图如下所示：



而对于超图的一个严格的数学定义，维基百科上是这样写的：

In mathematics, a hypergraph is a generalization of a graph, where an edge can connect any number of vertices. Formally, a hypergraph H is a pair H = (X,E) where X is a set of elements, called nodes or vertices, and E is a set of non-empty subsets of X called hyperedges or links.

k-均匀超图（k-uniform hypergraph）

对于超图而言，还有一个k-均匀超图的概念（k-uniform hypergraph）。它指超图的每个边连接的顶点个数都是相同的，即为个数k。所以2-均匀超图就是我们传统意义上的图，3-均匀超图就是一个三元组的集合，以此类推。

While graph edges are pairs of nodes, hyperedges are arbitrary sets of nodes, and can therefore contain an arbitrary number of nodes. However, it is often useful to study hypergraphs where all hyperedges have the same cardinality: a k-uniform hypergraph is a hypergraph such that all its hyperedges have size k. (In other words, it is a collection of sets of size k.) So a 2-uniform hypergraph is a graph, a 3-uniform hypergraph is a collection of triples, and so on.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

图6.4：超图和对应的两偶图

比如，在社会网络中，我们想在社区里面划分出家庭，但是家庭不是仅仅由两个人组成，这时候我们就用到了超图，超图的一条边可以连接两个以上的顶点，这样的边称为超边。



表6.2：超图和双向图。可以被表示为超图的例子。

本书中很多网络都可以被表示为超图。在社会学中，这样的网络被称为“关系网络affiliation networks”，并且我们在3.5节中见到了不少这样的例子。公司董事会的决策者们，合著文章的科学家们，在同一部电影的演员们都是这样网络的例子（见表6.2）

但我们不详细讨论这种网络，因为它可以被另一种方便的网络代替——二分网络/对分网络/二部图（bipartite network）。

### 6.6 二分网络

超图往往可以用二分网络表示，在社会学文献中被称为two-mode network。在这样的网络中存在两种结点，一种代表原来的结点，另一种代表着原来结点所属的组。我们在3.5节讨论关系网络affiliation networks讨论过二分网络，以及4.3.2节讨论推荐网络时。比如……一些例子……。

还是个例子。

二分网络的邻接矩阵叫做incidence matrix关联矩阵。如果n是人数，g是组数，那么the incidence matrix B is a g × n matrix

 (6.15)

比如图6.4b的incidence matrix是4X5的 (6.16)

尽管二分网络可能给出了特定网络的最完整的表达，但一般来说还是直接在普通的网络上做比较方便。我们可以使用二分网络去推断这些连接，创造一个one-mode projection from the twomode bipartite form。例如，再次考虑电影和演员的例子。我们可以做这样的映射去构造一个n个结点的网络，如果两个演员出演同一部电影，那么两个结点间有连接。对应的one-mode projection对于电影来说是一个g个结点的网络，如果两部电影都使用了同一个演员则两个结点间有连接。图6.5展示了这一方法。

图6.5：一个二分网络的两个one-mode projections

当我们形成一个one-mode projection，在二分网络中的每一个组导致了顶点的簇，在one-mode projection中他们相互连接，用网络的术语叫做“clique”团、派系（详见7.8.1节）。比如，在一个二分网络中如果一个组包括4个成员，那么这四个在one-mode projection中是相互连接的，如图6.5所示。所以projection概括来说，是cliques的集合。

one-mode projection被广泛地使用，但是它的结构丢弃了一些原来bipartite network的结构，因此，在某种意义上，它是一种不太强力的数据表现形式。比如，projection损失了两个结点共享组的数量的信息。比如在演员和电影的例子里，有些演员一起参演过很多电影，这两个结点间就会有很强的连接。

我们可以将这些信息用权重表示，用权值表示两个结点共享组的数目。但是，这样的权重网络还是不能捕获得原来的bipartite的全部信息——它不能记录组数，但也是种改进并被广泛使用。

数学上the projection可以用关联矩阵表示，如下B所示。*BkiBkj*=1当且仅当i和j都属于bipartite network中同一组k。因此用Pij表示i和j同属组的数量，(6.17)

Bki只能取0或者1，所以……

这样我们就可以通过计算P = BTB来给one-mode projection赋权值。

另一个one-mode projection，对于组的映射，可以用一个gXg的矩阵Pʹ = BBT来表示，非主对角线元素表示了两个组i和j共有的成员数，主对角线元素表示了一个组的成员数。

另一种二分网络的情况是有向的，比如5.1.1节的代谢网络——见图5.1a。在这个例子中更复杂的都有，但是他们的使用非常罕见，我们也不花时间在这上面。

### 6.7 树



图6.6：同一棵树的；两种草图。

书是连接的，无向网络，没有闭环——见图6.6a。连接的表示没有孤立的点，没有环的就是一棵树，很多棵树就是森林。

……一些对树的介绍……，对于顶点的选择，有时候有特点的原因。比如dendrogram系统树图[生物]。

现实生活中的网络树并不多，河流网络是一个天然形成的树结构的例子（见图2.6）。在第12章的时候会我们会学到“随机图”的模型，在这个模型中local groups of vertices—the so-called small components in the network形成树，并且我们探索了他许多特性……。在11.11.1节我们介绍了“dendrogram”，