# 互联网基础

作业讲解

# 第一讲作业:

2、简要说明目前互联网的层次化组成结构。

主要分为三个层次,核心骨干网(实现远程的互联互通)、城域网(实现区域范围的覆盖)、局域网(将各种终端设备接入到互联网)。

# 3、交换网络的主要优缺点是什么?

交换网络通过中间节点中继和转发信息;特点是资源复用,一份资源被多个用户同时共同使用。优点:利用率高,通信成本分担;缺点:时延增大。

# 4、电路交换与分组交换各自的主要优缺点是什么?

电路交换最主要的优点是可以提供高质量的传输保证,缺点是链路的利用率低,通信成本较高。分组交换最主要的优点是链路利用率高,通信成本低,缺点是传输质量保证比电路交换方式低。

# 5、数据发送和接收中, 封装和解封装的作用是什么?

从层次化的网络体系结构看,网络数据传输的发送过程也是数据封装的过程,将本层控制和管理用的信息加载在头部;而网络数据传输的接收过程也是数据解封装的过程,根据提取的控制和管理信息,对数据做相应的处理。

6、简要说明IP层和TCP层完成的主要任务,以及该模型的开放性和兼容性。

IP层的协议及其相关协议主要承担:按标准格式对数据包封装和解封;数据包的转发;以及转发包所需路由表的生成和维护等任务。TCP层有关协议主要承担:为应用程序提供标准化网络传输的调用接口;可靠的端到端数据传输等任务。沙漏模型的"开放性"就是在沙漏的中间层实现全球范围的标准化,兼容性就是中间层之外的标准,以及硬件和软件系统完全放开,允许各个组织、厂家或个人参与设计与开发。

7、写出你的计算机配置的IP地址、地址掩码、默认网 关、DNS,并且,简要说明它们的用途。

IP地址在互联网上唯一地标识一台计算机或者终端设备。默认网关负责接收和转发目的地址不属于自己IP子网的分组。DNS将文字形式表示的网络地址转换为二进制数字表示的IP地址。

- 8、为什么IPv4会出现地址危机问题?缓解这种地址危机的方法有哪些?
- 1) 利用可变长子网掩码来进一步划分出多个网段; 2) 利用网络地址转换NAT技术减少本单位对IP地址的需求; 3) 使用IPv6。

- 9、IPv6地址的长度是多少?接口标识符的用途是什么?128位。接口标识符:自动实现IP与MAC地址的绑定(映射各种类型的LAN技术,比如以太网及以后物联网时代可能出现的新的地址.)
- 10、外部路由协议和内部路由协议有什么不同用途?
- 11、单播路由与组播路由有何相同与不同?
- 12、为什么说"TCP是可靠的,而UDP是不可靠的"?

在TCP协议中使用了接收确认和重传机制,而UDP是尽力传送。

# 13、解决TCP对网络流量调控能力不足的方法有哪些?

- 1)建立基于多路径路由,进行流量分流,实现流量分布的负载均衡;2)是要在网络的所有中间节点加入检测、调节和控制功能,如分类、排队、调度等机制。
- 14、列举一些你经常使用的互联网应用程序。
- 15、为什么说Web网是"网络之上的网络"?

一方面,它利用互联网技术实现信息(网页)在物理网络中的存储和传播;另一方面,其信息网页的超链接又将这些网页组成了一个独立于底层物理网络的复杂网状结构的信息网络系统。

# 16、FTP、SMTP、HTTP等都采用TCP协议实现数据传输,

## 这种做法的好处是什么? 缺点是什么?

使用TCP的最大好处是端到端传输的可靠性有保证,可实现数据的无差错、 无丢失、无失序、无重复通信。吞吐量上不去。

# 17、"云计算"采用的的网络应用模式是什么? 其主要优点是什么?关键技术有哪些?

"云计算"采用集中式的客户/服务器模式。其最大优点是用户不再需要自己建设昂贵的服务器,避免重复投资,设备使用效率高,集中管理服务器有利于服务器端的系统优化和节能。

- 18、Web网页的地址信息如何标识?
- 19、搜索引擎的主要工作流程是什么(自己总结)?

数据搜索和获取;对获取信息的加工处理;提供相关信息服务等 三个主要流程。

20、Web服务系统的组件和相关协议有哪些?它们各自 承担的主要任务是什么?

WSDL: 描述服务; UDDI: 发布、查询服务;

SOAP: 执行服务; WSFL: 将分散的、功能单一的服务组织成一个复杂的应用

21、数据加密可以避免那些安全问题?

数据加密可以避免信息泄露,另外结合身份鉴别技术可以防止非

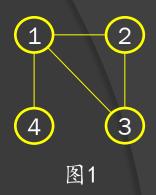
法登录。

- 22、什么是对称密钥和非对称密钥?
- 23、在网上购物时,你需要进行那些身份认证操作? 简要分析一下这样处理的好处是什么?
- 24、从网络组成布局看,在何处布设防火墙效果最好? 将防火墙布置在接入网与汇聚网或者核心网连接的网关处效果最好。一旦分组进入核心网络(汇聚网、骨干网),就很难再跟踪、捕获了。

# 第三讲作业:

- 25、1) 写出图1的邻接矩阵;
  - 2) 写出基于邻接矩阵计算节点度值 $k_i$ 的 向量K的表达式:





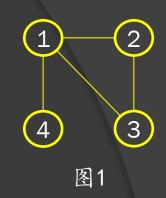
$$(1) \ A = \begin{pmatrix} 0 \ 1 \ 1 \ 1 \\ 1 \ 0 \ 1 \ 0 \\ 1 \ 1 \ 0 \ 0 \end{pmatrix}$$

(2) 
$$\mathbf{A} \cdot \mathbf{1} = \mathbf{K}$$
,  $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = (3 & 2 & 2 & 1) = (k_1, k_2, k_3, k_4).$ 

(3) 边数 
$$m = \frac{1}{2} \sum_{i=1}^{4} k_i = \frac{1}{2} (3 + 2 + 2 + 1) = 4$$
 (条)。

#### 26、写出图1中的最小割端集和最小割边集。

最小割端集为 $\{v_1\}$ 。注意:点2、3、4都不是,因为删除了其中的某单个点后并不影响其他点之间的连通性。最小割边集 $\{e_{14}\}$ 



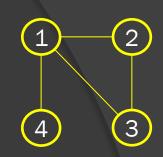
#### 27、利用连通图的数学定义证明图1是连通图,写出计算过程。

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A^{1} = A, \quad [A^{2}] = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \quad [A^{3}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

连通网络的充分必要条件:  $I + A + A^2 + A^3$ 为正矩阵, 前面的矩阵代入可得

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 2 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \end{pmatrix}.$$

#### 28、分析图1中哪个点的接近中心性最好?写出分析过程。



节点1的"接近中心性"最好。

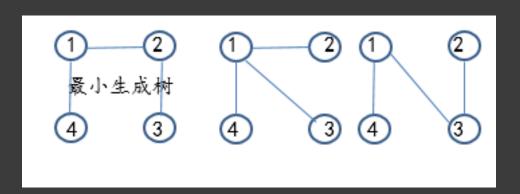
因为,
$$C_i = \frac{n-1}{\sum_{j \neq i} \overline{l_i}}$$
。 $C_1 = \frac{4-1}{l_{12} + l_{13} + l_{14}} = \frac{3}{1+1+1} = 1$ , $C_2 = \frac{3}{1+1+2} = \frac{3}{4}$ ,图1

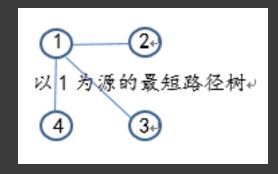
$$C_3 = \frac{3}{1+1+2},$$

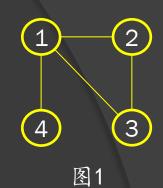
$$C_4 = \frac{3}{1+2+2} = \frac{3}{5}$$
.  $C_1$  最大。

最小边割集为e14。

29、图1有几棵最小生成树?举一例说明图1的最小生成树和最短路径树可以是同一棵树。

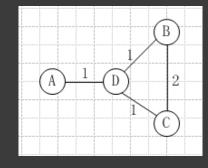




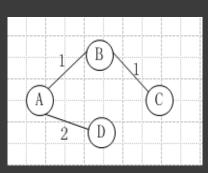


- 30、1)写出图2对应的共引矩阵和共引网络, 以及文献耦合矩阵和文献耦合网络;
  - 2) 写出图2的google矩阵Â;

(1) 共引矩阵和网络
$$AA^T = \begin{bmatrix} 2001\\0221\\0221\\1112 \end{bmatrix}$$



文献耦合矩阵和网络 $A^TA=\begin{bmatrix} 3 & 1 & 0 & 2 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$ 



(2)google矩阵

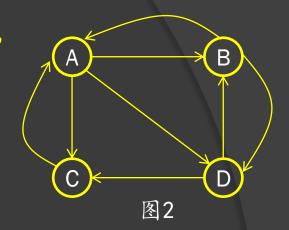
$$\begin{bmatrix} 2 & 2 \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\hat{a}_{ij} = \begin{bmatrix} 1/k_j^{out}, & \mathbf{f}_{j}$$
指向 $i$ 的边;  $0, \qquad$ 其它。

Α

图2

# 30、3)写出PR(1)的计算过程(α = 0.85),迭代过程大约多少步可以收敛?



$$PR(1) = \alpha \hat{A}^T PR(0) + (1 - \alpha) \frac{1}{n},$$

$$PR(1) = 0.85 \times \begin{vmatrix} 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 1 & 0 \end{vmatrix} \begin{vmatrix} 1/4 \\ 1/4 \\ 1/4 \end{vmatrix} + 0.15 \begin{vmatrix} 1/4 \\ 1/4 \\ 1/4 \end{vmatrix}.$$

- 31、1) 说明在图3中标记出一个"3-核"的生成过程;
  - 2) 分别计算两个实心点的聚类系数;
  - 3) 计算两个实心点的余弦相似性;



4) 采用皮尔逊相关系数的优点是什么?

(1) 首先,删除度为1的节点,重复该过程,直到再找不出度为1的节点为止;其次,删除剩余图中度<=2的节点,重复该过程,直到再找不出度为2的节点为止。至此,剩余图中的节点的度数都不小于3。

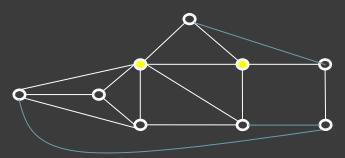
(2) 聚类系数分别为: 
$$C_i = \frac{2E_i}{k_i(k_i-1)}$$
, 6/15=0.4 ; 1/3=0.333

(3) 余弦相似性 
$$\sigma_{ij} = \frac{n_{ij}}{\sqrt{k_i k_j}} = \frac{2}{\sqrt{24}} \approx 0.41$$

(4) 皮尔逊相关系数能更好地表现两者之间是否具有共同特质。如果两者都较多地高于均值或者都低于均值,说明他们特质相近,其皮尔逊系数为较高的正值;如果两者都接近均值,则说明两者都比较平常(无突出特质),其皮尔逊系数接近0值;如果一个高于均值,另一个低于均值,说明他们特质正相反,其皮尔逊系数为较低的负值。

- 32、1) 说明连通度与可靠度的相同与不同之处;
  - 2) 如何使图3的连通度不小于3? 画图说明。

- (1) 连通度是对物理连接拓扑关系的评估。可靠度是在物理连接拓扑之上,还需要对点和边正常工作的程度(概率)进行评估。 K-终端可靠度:在一个网络的边连接概率图中,确保指定的K-个节点构成一个连通图的概率p,记为 $Rel_K(G)$ 。 连通度:= min {最小割点集的点数,最小割边集的边数}
- (2) 增加连边, 使得每个点的度数不小于3, 并且网络不存在只有2个点的最小割端集。

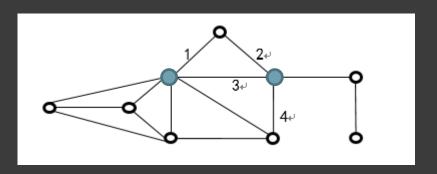


# 32、3) 空心点与边可用概率为p,写出实心点间 Rel2 计算过程。

两个实心点之间的最小割边集为<2,3,4>和<1,3,4>; 两个实心点不连通情况为<2,3,4>都不通,或者<1,3,4>都 不通;

<2,3,4>都不通的概率为 $(1-p)^3$ , <1,3,4>都不通的概率为 $(1-p)^3$ ;

两个实心点不连通的总概率为 $2(1-p)^3$ ; 两个实心点的2-终端可靠度 $Rel_2 = 1 - 2(1-p)^3$ 。



# 第四讲作业:

33、求背包问题的最优解属于NPC问题, 其蛮力算法如下:

输入: 背包最大承重量c, 物品个数n, 单个物品重量w[n], 价值v[n]。

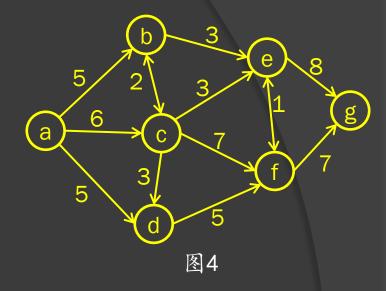
输出:装在背包的物品以及产生的最大价值。

- 1. 初始化最大价值Vmax=0, 结果子集s=φ;
- 2. 循环, 对集合 {1, 2, ···, n} 的每个排列T情况, 执行:
  - 2.1 初始化背包的价值v=0, 背包的重量w=0;
  - 2.2 循环,对T的每个元素j:
    - 2. 2. 1 如果w+w<sub>j</sub><c,则w=w+w<sub>j</sub>、v=v+v<sub>j</sub>;
    - 2.2.2 否则, 跳出循环, 转步骤2.3;
  - 2.3 如果V<sub>max</sub><v,则V<sub>max</sub>=v、s=T,转步骤2;
- 3. 输出s中的元素。

写出该算法计算复杂度的数学分析表达式。

所有可能排列情况,共有 $n! \approx n^n$  循环之中可能需要遍历所有排列的成员,量级为n。所以,整个算法的复杂度为 $O(n^n * n) \sim O(n!)$ 

34、图4中,边的权值代表链路容量, 以源节点到目的节点的路径带宽最大化 为优化目标。基于Dijkstra算法计算以 a为源到所有节点的的最大带宽通路树, 写出完整的计算过程。



解法1:修改的Dijkstra算法

#### Algorithm · Dijkstra

Input: a network G, a source node s and a destination node t

Output: a path from s to t with the maximum bandwidth

- 1. for each node v in G do  $\{P[v] = 0; B[v] = -\infty\}$ ;
- 3. for each neighbor w of s do  $\{P[w] = \sum_{s} B[w] \stackrel{\text{EB}}{=} b([s, w]); \text{ add } w \text{ to } F; \}$
- 4. repeat

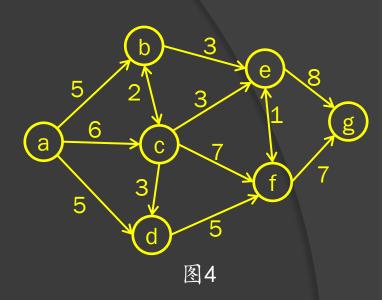
remove the node u of maximum B[u] from F

for each neighbor w of u do

case 1. 
$$B[w] = -\infty$$
. 尚未搜索到的点  $\{P[w] = u; B[w] = \min \{B[u], b([u, w])\}; add w to F\}$  case 2. (w is in F) and ( $B[w] \le \min \{B[u]; b([u, w])\}$ ):  $\{P[w] = u; B[w] = \min \{B[u]; b([u; w])\}; \}$ 

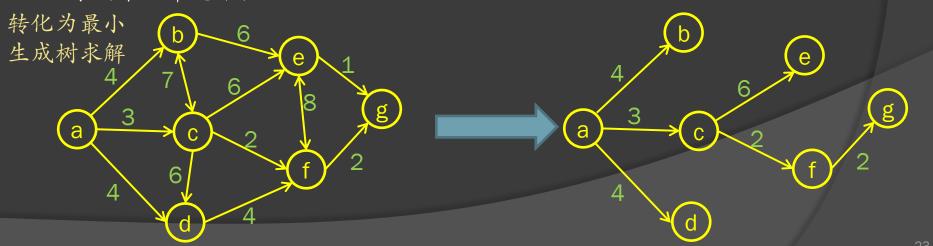
until  $B[t] \neq -\infty$  and t is not in F

34、图4中,边的权值代表链路容量, 以源节点到目的节点的路径带宽最大化 为优化目标。基于Dijkstra算法计算以 a为源到所有节点的的最大带宽通路树, 写出完整的计算过程。

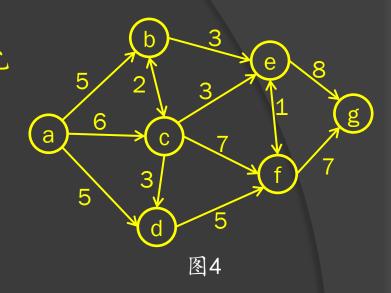


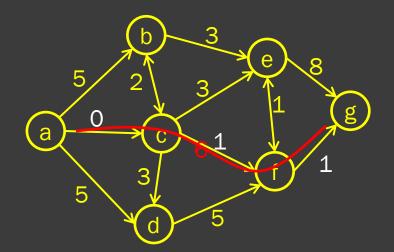
#### 解法2: 基于最大生成树的解法

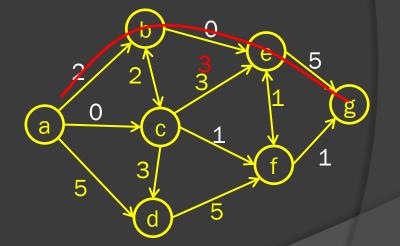
可以证明:设G为一网络,C中的边e的带宽值为b(e), CT是G根据链路带宽的最大生成树.那么对于G中的任意两个结点s和切T中从s到t的唯一的路径Pst就是G中从s到t的最大带宽路径.



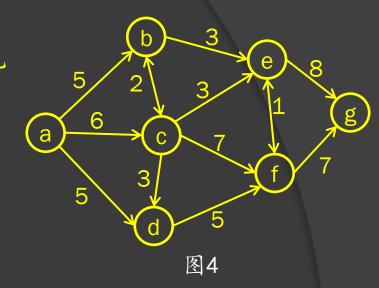
35、上题图4中的箭头代表着链路上允许数据流动的方向,基于Ford-Fulkerson计算点a到点g的最大流。

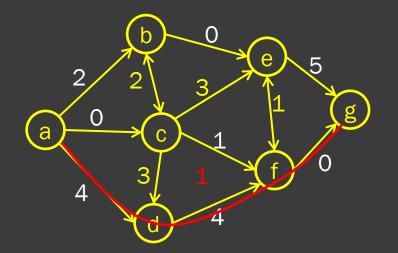


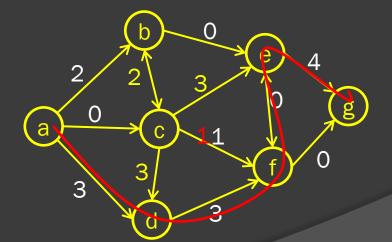




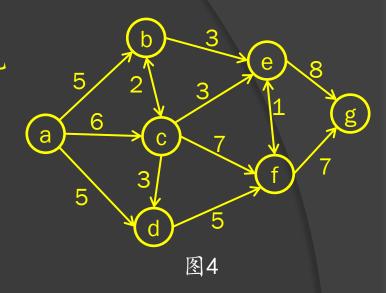
35、上题图4中的箭头代表着链路上允许数据流动的方向,基于Ford-Fulkerson计算点a到点g的最大流。

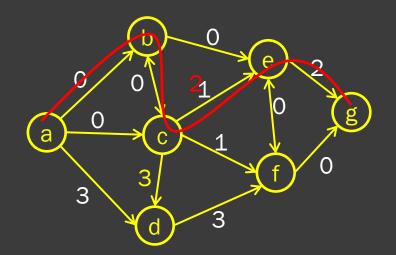


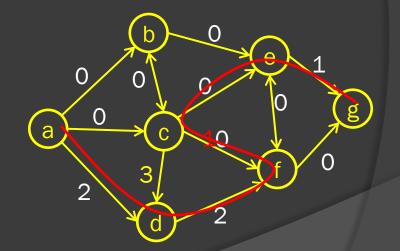




35、上题图4中的箭头代表着链路上允许数据流动的方向,基于Ford-Fulkerson计算点a到点g的最大流。







最大流14

# 36、基于Ford-Fulkerson方法(增广路径)来进行不相交路径的搜索的主要好处是什么?

简单采用最短路径策略可能会遇到实际路径存在,但却找不出不相交路径的情况。而 增广路径方式可以避免这种情况发生。

37、在图4中选出2条点a到点g的并发路径,用以分流,达到负载均衡的目的。说说你的选择的理由。

从流量最大化角度当然是a-c-f-g和a-b-e-g比较好。如果分流的流量小于3,可以选择a-d-f-g和a-b-e-g,这是从"并发路径相距越远越好"的思路来考虑的。另外,为了对c点实施保护,也可以用上面的两条路径来分流,实现网络流量工程中的流量保护。

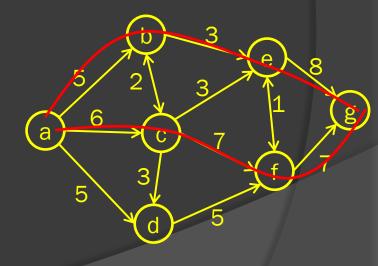


图4

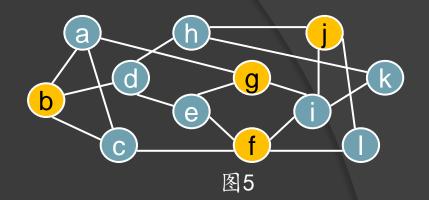
### 38、Kruskal算法和Prim算法适用的条件有什么不同?

前者适用于点数大于边数(我们实际的网络很多是这样)的情况,因为其计算复杂度为 $O(m \log n)$ ;后者正相反,其算法复杂度 $O(n \log m)$ 。

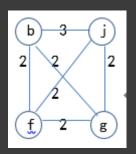
- 39、贪婪算法的特点是"以单步搜索的最优去逼近全局最优"。 参照"贪婪"思路, PPT课件中关于平面Steiner最小树的启发式算法应做怎样的修改?修改后的算法复杂度是多少? 怎样可以使选点的选取范围更小一些呢?
  - 1) 需要将启发式方法中"S集中的点数达到n-2"的判定条件删除。
  - 2) 计算复杂度在0 (mn)量级。
  - 3) 如果利用"凸包形状"和"两点之间的重边"进行联合约束,可较多地压缩候选点的个数。可以把重边端点作为s点的候选,重边端点是两条直角折线的重合点。例如,右图中a到b的直角折线与c到b的直角折线的交叉点f。

将随机选取点的过程改为每次选候选点中最大价值的那个点。

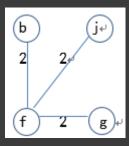
40、图5是一通信网络,其中橙色 节点是一组播组。基于KMB算 法计算该组的Steiner路由树。



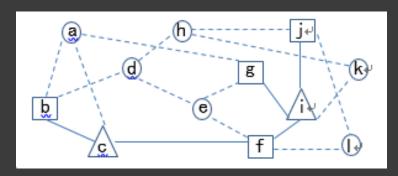
首先, 构造完全图



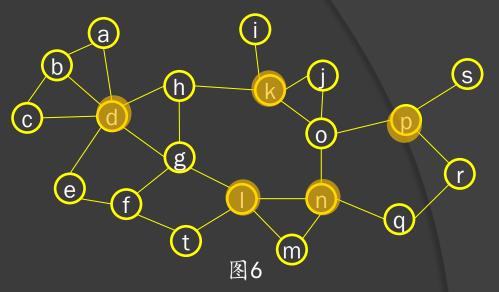
计算得到最小生成树



构造子图Gs; 计算Gs的最小生成树; 删除多余的叶子和边。



41、无线传感网常常采用组簇 方式的层次路由机制。普 通传感节点将检测数据以 单跳近距离通信方式传给 簇头,簇头再将数据中继



给连接互联网的网关。所以,簇头既具备传感节点的功能, 又具备远距离通信的路由功能,但其价格要比普通传感设备 贵很多。图6是一传感网的一跳邻居互连拓扑图,怎样部署 路由节点既满足应用需要,又使建设成本最低。

采用最小支配集方法来做。选择{d、n、k、p、t}五个点配置路由设备即可。先选择度数最大的节点d,去除关联的点和边;再在剩余途中选择度数最大的节点n;然后选出p;再后选出k;最后选t。

42、地图着色问题(相邻区域不同色) 可以借助图论中的一些方法来解 决。结合图7给出的地图,谈一 谈你的解题思路。用了几种颜色?

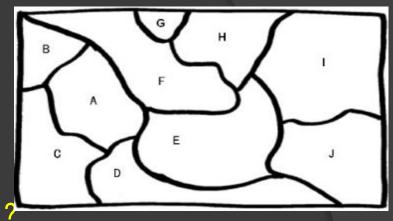
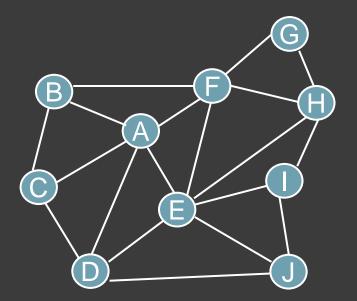
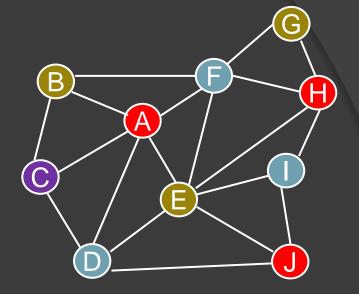


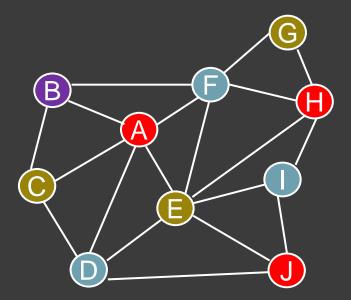
图7

建模成极大独立集问题 (独立集的点之间都不是邻居关系)

- (1) 画出映射图来,区域用点表示,相邻区域的两点之间连一条边。
- (2) 要求同色节点集合中两两不相邻,所以,同色点实际是一个独立集。先找出图的一个极大独立集,分配一种颜色;然后找出另一个极大独立集,在分配一种颜色;直到所有点都被着色。
- (3) 因为所有点都需要着色,所以,选求出所有极大独立集,然后找出极大独立集的并集可以包含所有节点的那几个极大独立集的组合。
- (4) 在其中找出最少个数极大独立集的组合, 个数就是所需的最少颜色数。



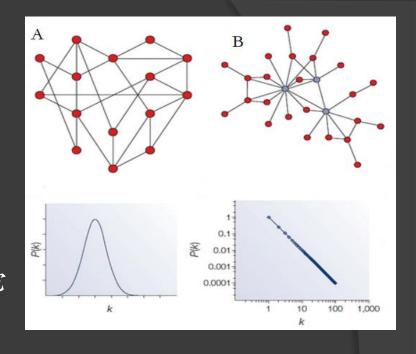




- 43、"图划分"与"社团发现"有什么不同? Kernighan-Lin算法的思路是怎样的?与其相比,试论述"谱方法"中采用了怎样的方法来降低计算复杂度?
  - 网络划分(图划分):将网络节点划分为指定规模、数量的非重叠群组,并使得群组之间互连的边数最少。
  - 社团发现: 找到网络内部不同群组之间的自然分割线。
  - "谱方法" > 特征值和特征向量 > 降维
- 44、你在网上搜索一下,最新的关于社团发现的研究有哪些?

# 45、图8是两种模型产生的网络拓扑和度分布情况, A和B分别对应哪种类型的网络模型?

A对应随机网络,其度分布为泊松分布。 B对应无标度网络,其度分布为幂率分布。它 们都有比较短的平均路径长度,以及比较低 的聚类系数。



# 46、用小世界模型生成的网络拓扑会与图8. A和图8. B存在哪些差异?

小世界网络的拓扑图中,邻居之间连边的数量要远高于图A和图B,因为小世界模型的聚类系数上要明显高于随机模型和幂律模型。另外,度分布特性上,小世界模型与随机模型类似。平均路径长度上,小世界模型也与随机模型接近,但要大于幂律模型。

- 47、怎样在一个规则网络上构造"捷径",才能使网络系统具备 "小世界"特性?分析该方式的优缺点。
  - 一种是随机地断键重连;另一种是随机地加入一个新连接。后一种方式在数学描述上比较容易,前一种没有清晰的数学描述。
- 48、度分布为 $p_k = (1 e^{-\lambda})e^{-\lambda k}$ 的配置网络, $\lambda > 0$ 。分别写出度分布和余度分布的生成函数 $g_0(z)$ 和 $g_1(z)$ 。设网络规模为n,写出某一节点的两跳邻居个数的计算表达式。

生成函数
$$g(z) = p_0 + p_1 z + p_2 z^2 + \dots = \sum_{k=0}^{\infty} p_k z^k$$
。 对 $p_k = (1 - e^{-\lambda})e^{-\lambda k}$ ,有 $g(z) = (1 - e^{-\lambda})\sum_{k=0}^{\infty} (e^{-\lambda}z)^k = \frac{e^{\lambda}-1}{e^{\lambda}-z} = g_0(z)$ 。 因有 $g_1(z) = \sum_{k=0}^{\infty} q_k z^k$ ,借助余度定义可得 $g_1(z) = \frac{1}{k} \frac{dg_0}{dz} = \frac{g_0'(z)}{g_0'(1)}$ , $\bar{k} = g_0'(1)$ ,则 $g_1(z) = \left(\frac{e^{\lambda}-1}{e^{\lambda}-z}\right)^2$ 。 任意一节点两跳邻居数量的期望值 $c_2 = g_0'(1)g_1'(1)$ 。前一项就是均值, $g_1'(1) = \frac{1}{k}\sum_{k=0}^{\infty} k(k+1)p_{k+1} = \frac{\overline{k^2}-\bar{k}}{\bar{k}}$ 。 先计算指数分布的期望值和方差,也就是度分布的一阶矩和二阶矩,代入给出 $c_2$ 的计算表达式。

# 49、你认为BA模型会产生图9这种比较 奇异拓扑图形的原因是什么?

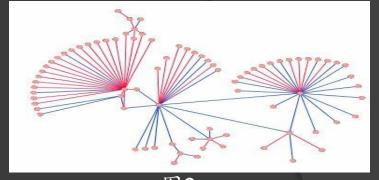


图9

两大原因:一是个别节点的度数极高;二是网络的聚类系数很低。这与BA模型的"优先连接"策略密切相关。优先连接包括了时间特性、适应独特性等,这也就是"赢者通吃"的效应。加上网络的平均距离很短,进一步降低了聚类特性,形成几乎是爆炸式星形的拓扑结构。