图论与代数结构

最大流与最小费用流

崔 勇
清华大学计算机系
网络技术研究所

第五章 匹配与网络流

- 网络流图
  - 掌握概念，学习玩命简化的建模思路
  - 掌握割切的概念和最大流定理
  - 学习算法设计的基本思路
- Ford-Fulkerson最大流标号算法
- 最小费用流
  - 掌握概念，了解基本求解思路
- 了解互联网中的网络流

网络流图

- 定义5.5.1
  - 一个运输网络N（或称网络流图）是一个没有自环的有向连通图，它满足：
    - 只有一个负度为0的结点s，称为源
    - 只有一个正度为0的结点t，称为汇
    - 每条边(i,j)都有一个非负实数权c_{ij}，称为该边的容量。如果结点i到j没有边，则$c_{ij}=0$.
- 网络流图：某种物品从产地s通过不同道路到销地t

简化：单一产品、单源单汇、无时间概念

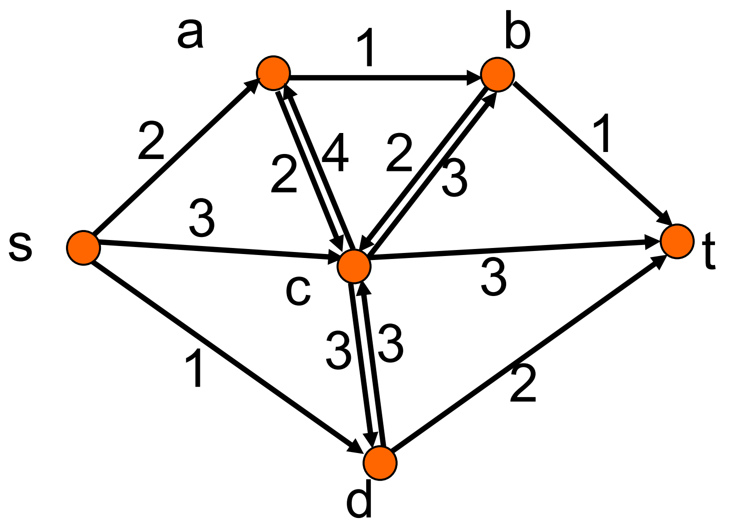
网络流图

- 在网络流图N中，每条边e_{ij}都给定实数f_{ij}，这一组f_{ij}称为该网络的允许流，如果f_{ij}满足：

  1. $f_{ij} <= c_{ij}$ (容量限制)
  2. 对固定的节点i，f_{ij} 之和等于 f_{ji} 之和(i不为s,t)

  流量守恒

  3. f_{sj} 之和等于 f_{jt} 之和, 记为w



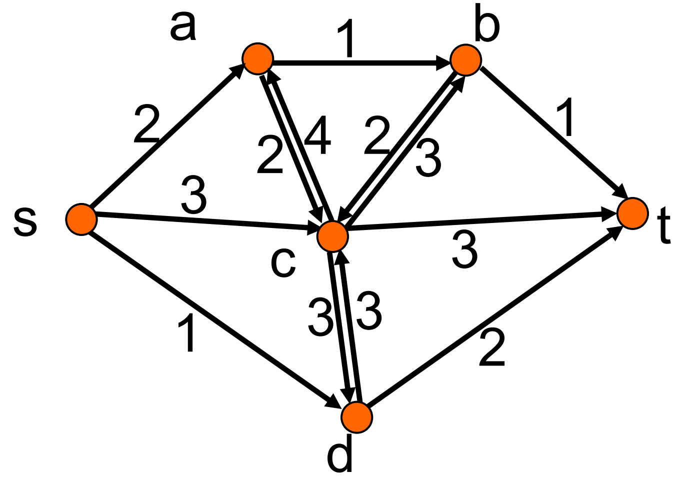
- w称为它的流量

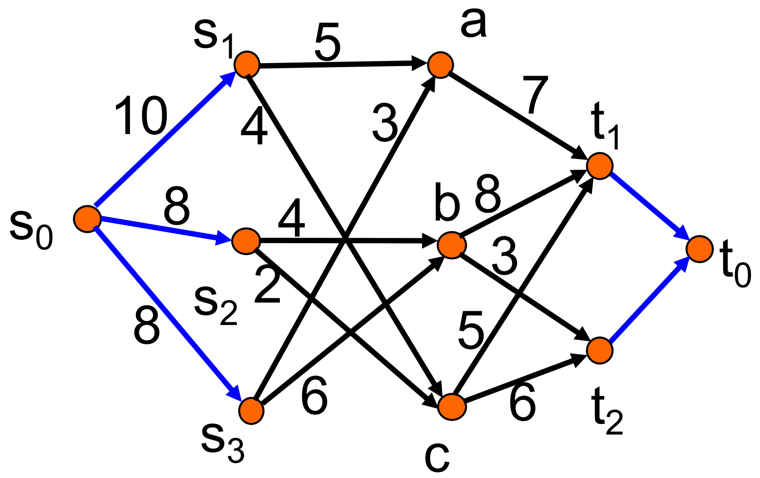
简化：容量有限，产品无限

基于特殊情况，继续给定义

网络流图

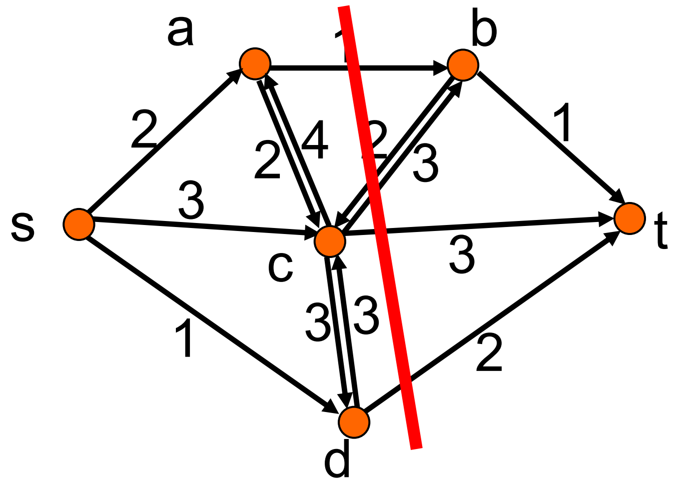
- 在网络N允许流分布f中，满足$f_{ij}=c_{ij}$的边称为饱和边，否则为非饱和边
- 网络的最大流
  - 某允许流分布使流量w最大：$w_0 = \max \sum_j f_{sj}$
- 如何处理多产地多销地的网络？
  - 增加超发点s_0和超收点t_0
  - 增加若干边(s_0,s_i)和(t_j,t_0)
  - 边(s_0,s_i) 的容量是s_i的生产能力
  - (t_j,t_0)容量是t_j的销售能力
  - 得到一个单源单汇的网络流图





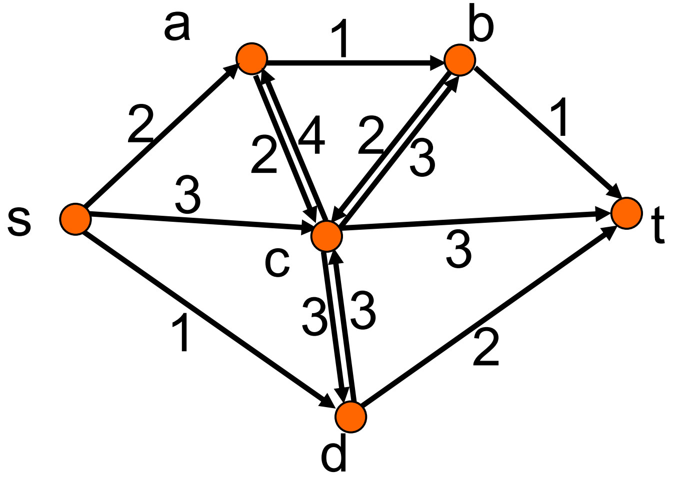
网络流图

- 定义5.5.2
  - 设S是网络流图N=(V,E) 中的一个结点集，满足
    1. s∈S
    2. t∈$\bar{S}$，$\bar{S}$ =V-S
  - 则全部有向边(i,j)，i∈S, j∈ $\bar{S}$ 的集合称为N的一个割切，记为$(S,\bar{S})$
  - $(S,\bar{S})$中各边的容量之和称为该割切的容量，记为$C(S,\bar{S})$



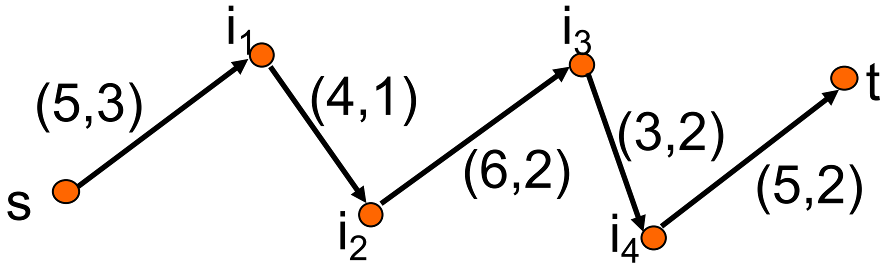
网络流图

- 例5.5.1
  - 令S={s}，则 $(S,\bar{S})$ ={(s,a),(s,c),(s,d)}， $C(S,\bar{S})$=6
  - 令S={s,a,c}，则$(S,\bar{S})$ ={(a,b),(c,b),(c,d),(c,t),(s,d)}，$C(S,\bar{S})$ =11
- 定理5.5.1
  - 网络的最大流量小于等于最小的割切容量，即max w≤min C
- 定理5.5.2
  - 网络流图N中，其最大流量等于最小割切的容量，即 $\max\omega=\min C(S,\bar{S})$



网络流图

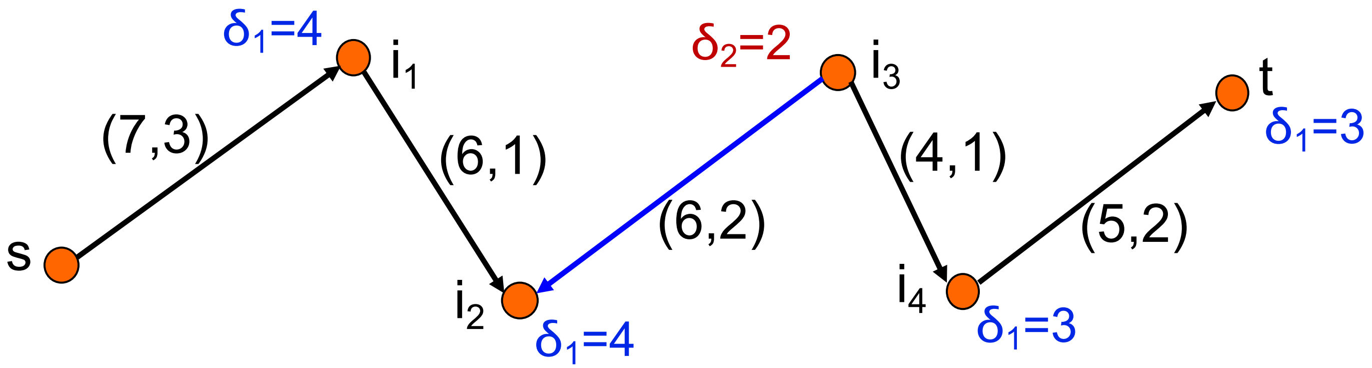
- 增流路径
  - 如果网络的允许流并不是最大流，就一定存在着从s到t的一条可以增加流量的路径，简称增流路径
  - 图中(a,b)，a表示其容量，b表示它当前的流
  - 令$s,i_1,i_2,\ldots,i_k,t$是一条从s到t的路径P_{st}
  - 其中若边的方向是从i_j到i_{j+1}，则称为前向边
  - 如果这条路径上每条边e_{ij}都有$f_{ij}<c_{ij}$，存在增流路径
  - 令 $\delta=\min\limits_{e_{ij}\in P_{st}}\left(c_{ij}-f_{ij}\right)$，使$P_{st}$每条边的流都增加δ
  - 结果仍然是网络的允许流分布，但流量比先前增加了δ



增流过程不影响其他边和节点

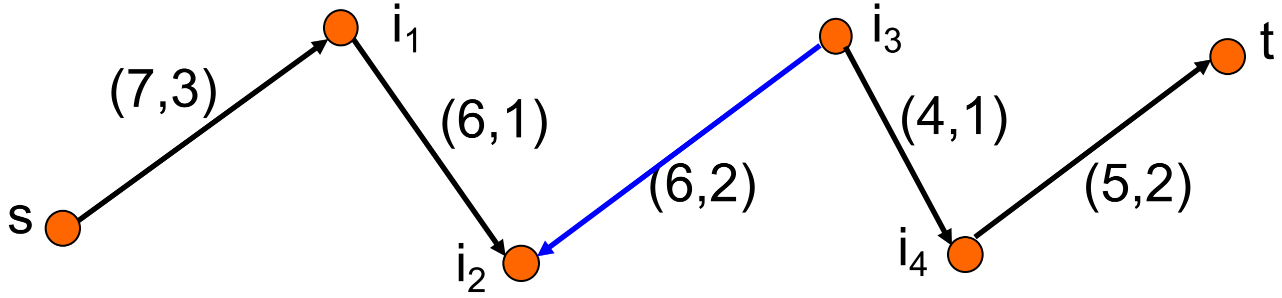
网络流图

- 具有后向边的增流过程
  - 汇点的流入量增加1是从i_4获得
  - i_4要保持流的守恒，应使f_{34}增加1
  - i_3的守恒是由i_3少供应i_2 1个单位流而得到保证
  - 增流路径中的后向边e_{ji}要求$f_{ji}>0$
  - i_2由于i_3少供应1，因此只有从i_1多索取1才能守恒
  - 本例增流瓶颈为蓝色后向边，增流量δ=2



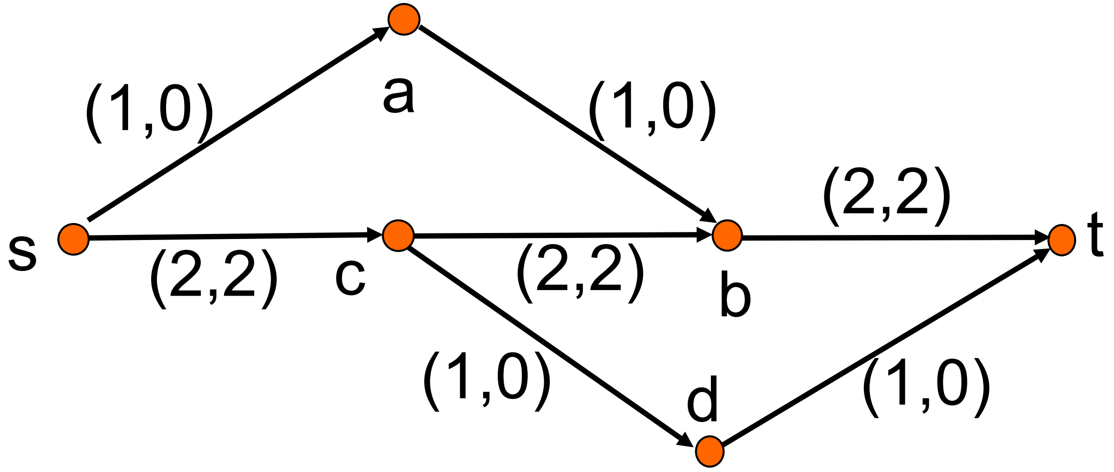
网络流图

- 在包含前向边和后向边的增流路径P_{st}中
  - 要求前向边e_{ij}满足$f_{ij}<c_{ij}$，向后边e_{ji}满足$f_{ji}>0$
  - 设P_{st}的全部前向边e_{ij}中，$\delta_1=\min(c_{ij}-f_{ij})$
  - 全部后向边e_{ji}中，$\delta_2=\min f_{ji}$
- 再令$\delta=\min(\delta_1, \delta_2)$,那么$P_{st}$中可增加流量$\delta$
- 在网络流图中只存在这两类增流路径



网络流图

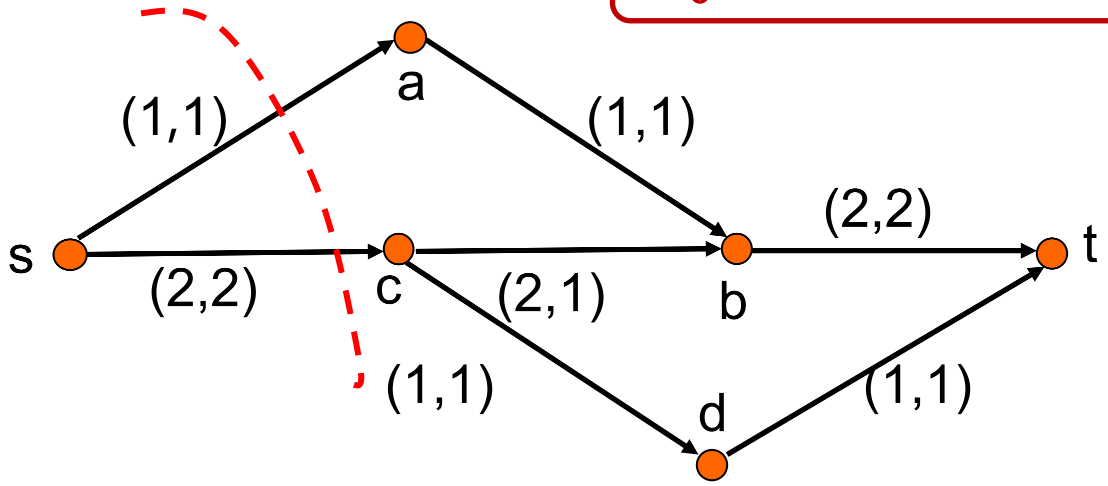
- 例5.5.2
  - 如果最初流量w=0，第一条增流路径如(s,c,b,t)
  - 它全部由前向边组成，δ=2，因此可增流2
  - 当前允许流分布：边(s,c),(c,b),(b,t)的流都是2，其余边均为0
  - 存在增流路(s,a,b,c,d,t)
  - 其中(c,b)是后向边，$f_{cb}=2$，其余都是前向边，满足$f_{ij}<c_{ij}$，这条路上δ=1



网络流图

- 例5.5.2 (续)
  - 因此增流之后得到下图，其中边(c,b)的流为1，这仍然是一个允许流分布
  - 此时网络中已不存在任何增流路径
  - 所以最大流量是$w_0=3$

$w_0=3$ 等于最小的割切容量

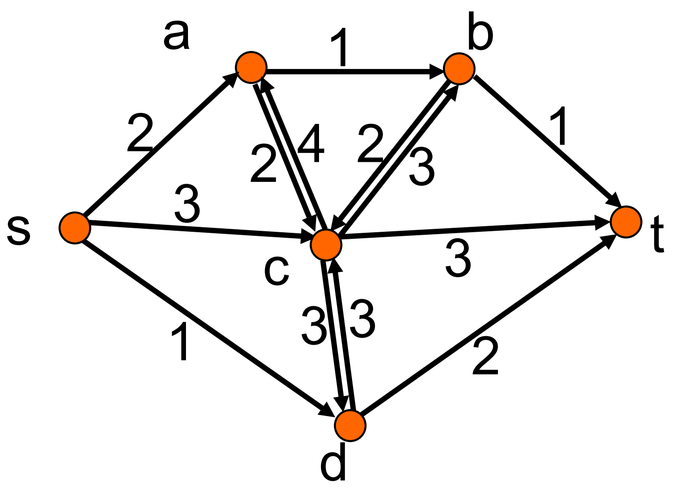


第五章 匹配与网络流

- 网络流图
- Ford-Fulkerson最大流标号算法
- 最小费用流
- 互联网中的网络流

Ford-Fulkerson 最大流标号法

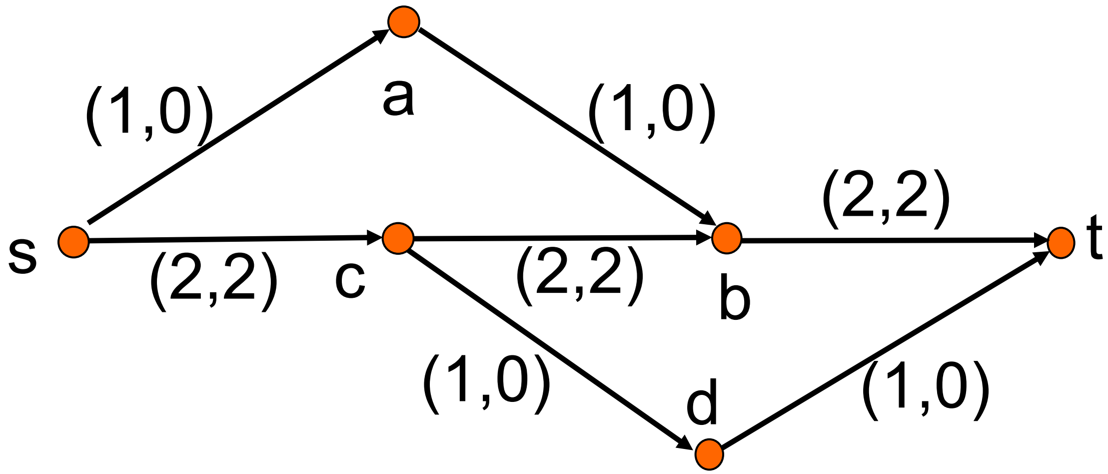
- Ford-Fulkerson最大标号算法简介
  - 以定理（最大流等于最小割切容量）为基础
  - (1) 标号过程（寻找增流路径的过程）
    - 从s向t标号，是否能标到t？
    - 如果不能标到t，则此时f是最大流，其流量为最大流
    - 否则在标号过程中最后能从结点s标到结点t，即找到s到t的增流路径，转过程(2)
  - (2) 增流过程
    - 沿着这条从s到t的增流路径增流，修正这条路上的流，得到新的允许流分布f',再转(1)



标号过程需明确增流路径和流量

Ford-Fulkerson 最大流标号法

- 结点v的标号(d_v, δ_v)
  - 标号过程中每个结点v都有一组标号(d_v, δ_v)
  - d_v 表示标号过程中结点v是因为哪个结点才得到标号的，称为标号来源结点
  - d_v也表示标号的方向（正向或反向）
  - 若v得到标号，表明网络里存在一条s到v的增流路径P，其最大增流量是δ_v

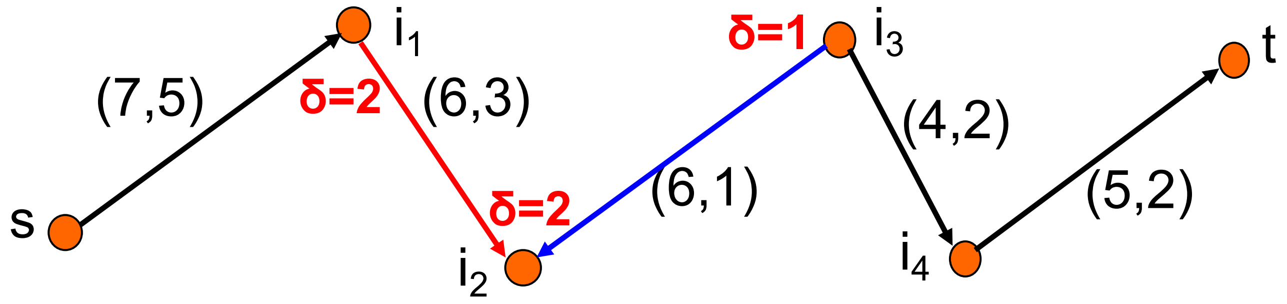


Ford-Fulkerson 最大流标号法

- 算法过程

流量变化仅影响当前路径

  - 首先对源s标以(-,∞),其中标号来源结点d_s的值为空



  - 设e是连接u和v的边，假定u已标号，而v尚未标号
  - 正向标号：若e=(u,v)且f(e)<c(e)，则标号方向为正，v的标号(u^+, δ_v)，其中$\delta_v = \min(\delta_u, c(e)-f(e))$
  - 反向标号：若e=(v, u)且f(e)>0, 则标号方向为负，v的标号(u^-, δ_v)，其中 $\delta_v = \min(\delta_u, f(e))$

Ford-Fulkerson 最大流标号法

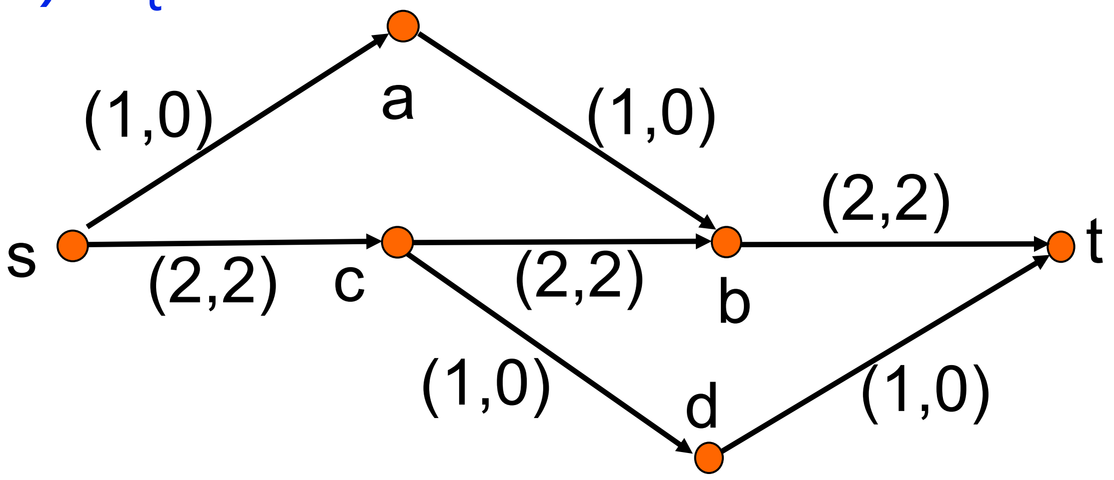
- 算法过程（续）
  - 在标号过程中，每个结点最多进行一次标号
  - 最终结点t或者能得到标号，或者无法得到标号
  - 若t得到标号，则由标号规则可确定一条s到t的增流路径P_{st}，它可以增流δ_t
  - 增流过程：回溯检索这条路径并修改标号，得到新的允许流分布f'

Ford-Fulkerson 最大流标号法

- Ford-Fulkerson算法描述如下
  - S1. 在给定的网络流图中任一流分布f（如令N中每条边e，都有f(e)=0）
  - S2. (标号过程开始）给s标号(-, ∞)
  - S3. 如果存在一个未标结点v，它可以通过正向标号或反向标号得到标号$(u^{+/-}, \delta_v)$，则标之并转S4，否则结束(此时f已是最大流分布)
  - S4. 如果v=t转S5，否则转S3

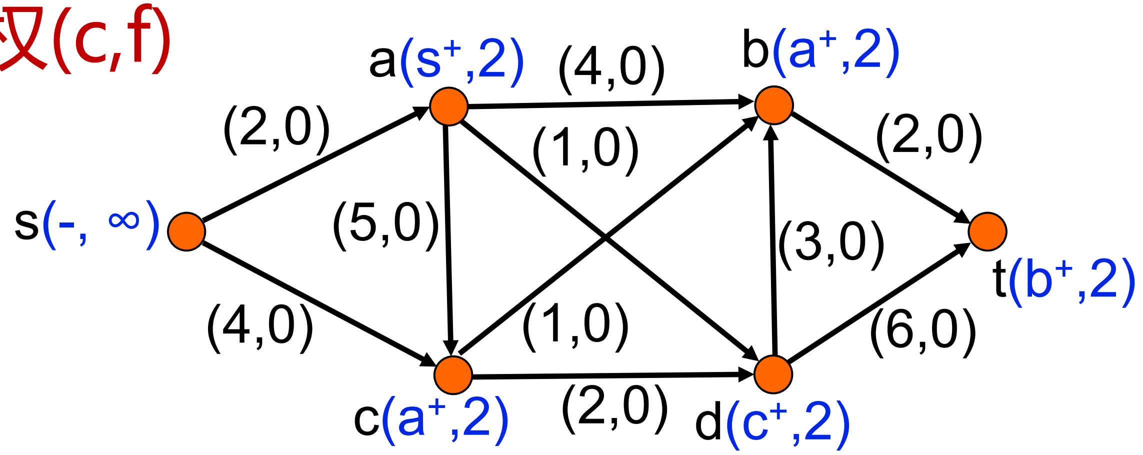
Ford-Fulkerson 最大流标号法

- 算法描述(续)
  - S5. (增流过程开始)，设v的标号是(d_v,δ_v)
    - 1.若$d_v=u^+$，则令$f(u,v)=f(u,v)+\delta_t$
    - 2.若$d_v=u^-$，则令$f(u,v)=f(u,v)-\delta_t$
  - S6. 若u=s，删去全部标号转S2，否则令v=u，转S5



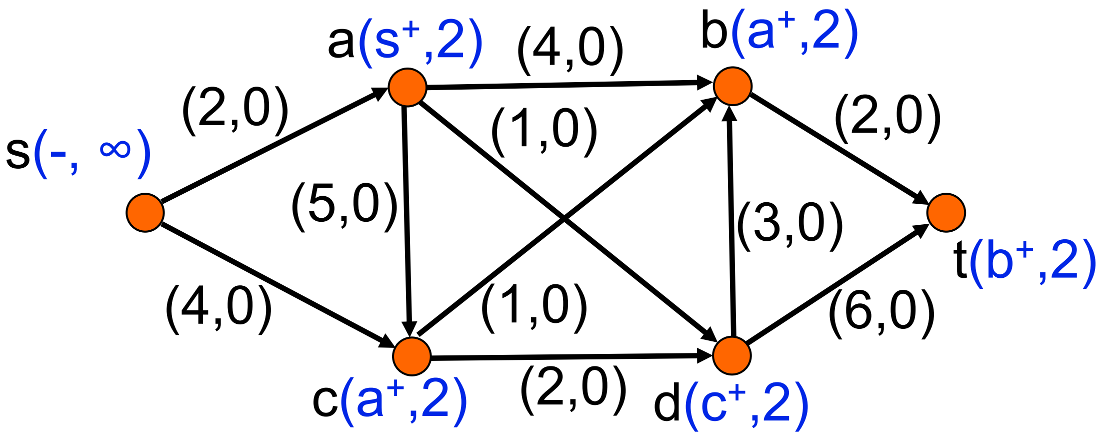
Ford-Fulkerson 最大流标号法

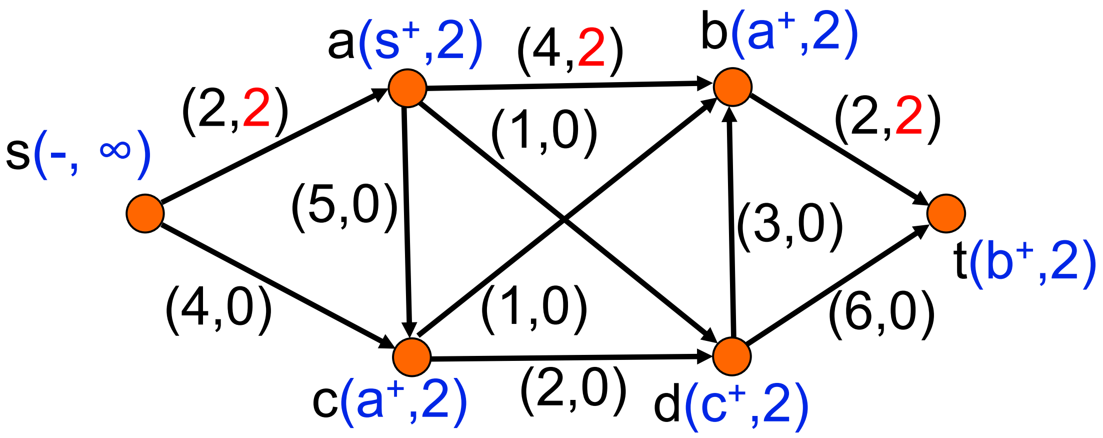
- 例5.6.1
  - 运输网络N的每条边e都有两个权(c,f)
  - 最初N中每条边有 f(e)=0, 即流量w=0
  - 开始标号时结点s为(-, ∞)
  - 依次给结点a,b,c,d和t的标号, 完成标号过程
  - 在增流过程中确定了增流路径P_{st}
  - 从后向前回溯，增流路径：s→a → b → t



Ford-Fulkerson 最大流标号法

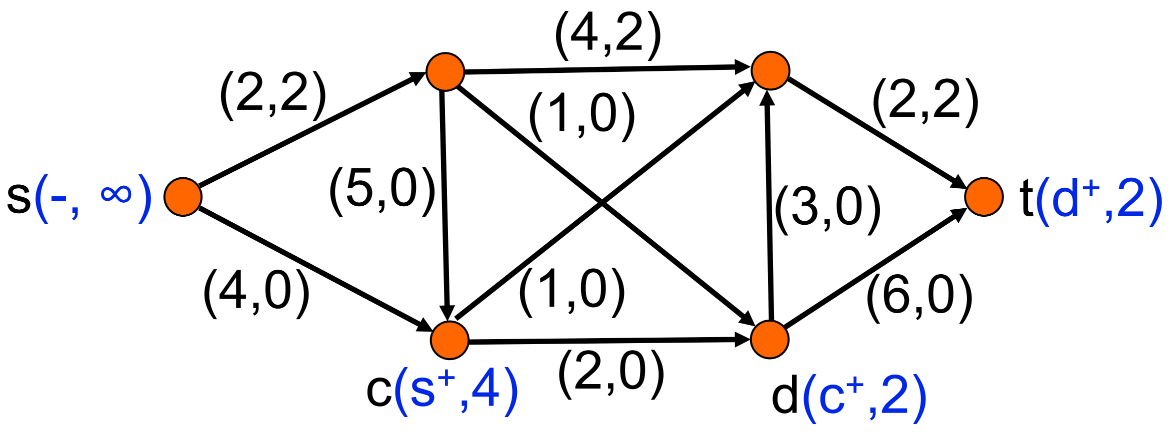
- 例5.6.1(续)
  - 标号均为正
  - 所有边都是前向边
  - 每条边可增流δ_t =2
  - 进行增流
  - 网络流量w=2
  - 是最大流了吗？



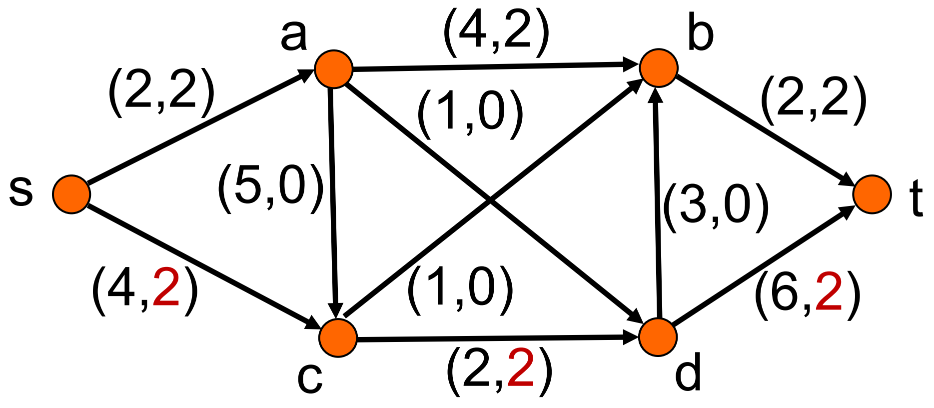


Ford-Fulkerson 最大流标号法

- 删去所有标号，从s开始重新标号

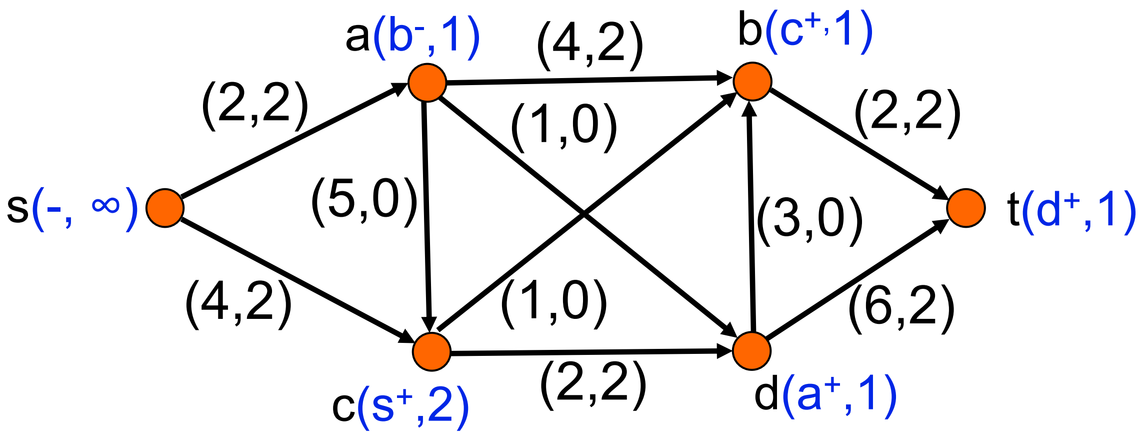


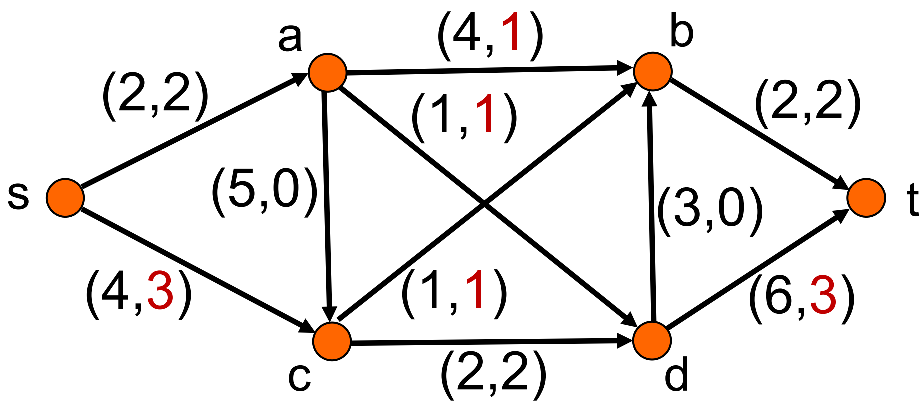
- 得到新增流路径(s,c,d,t),可增流δ_t=2，此时w=4



Ford-Fulkerson 最大流标号法

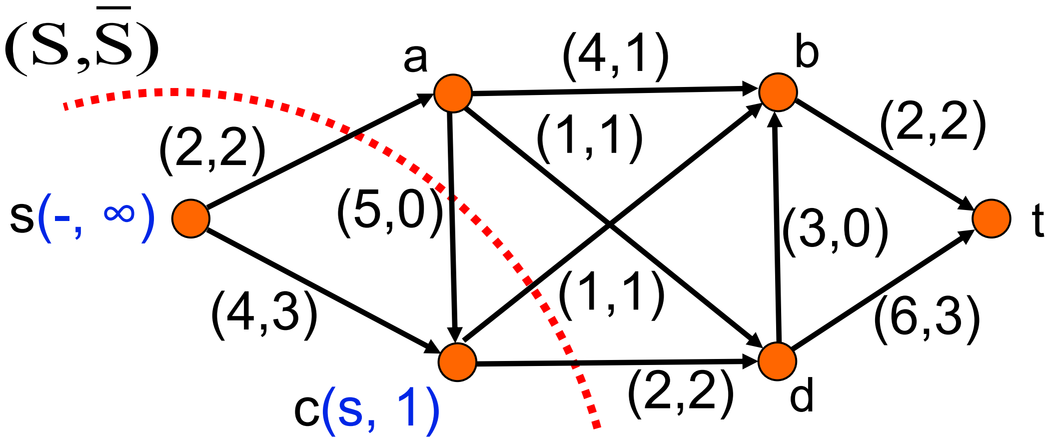
- 经过标号过程又得到一条增流路径(s,c,b,a,d,t)
- 此时边(a,b)是后向边，其余都是前向边
- 增流：这时向前边的流增1而后向边的流减1，总流量w=5





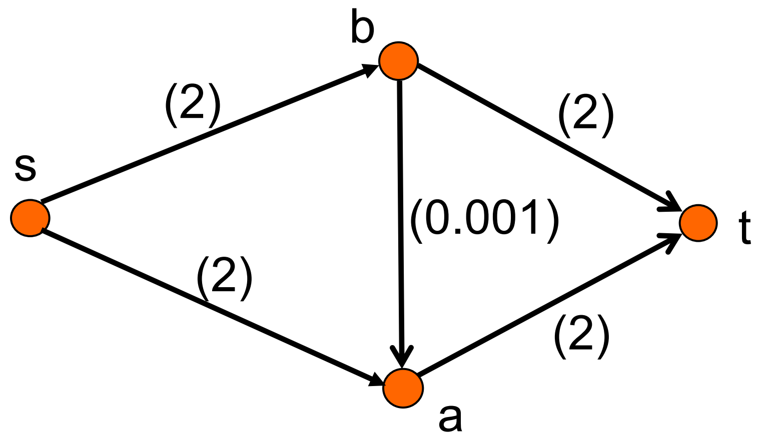
Ford-Fulkerson 最大流标号法

- 例5.6.1 (续)
  - 再从S开始标号，只能标到c，无法标到t
  - 因此不存在s到t的增流路，w=5是最大流
  - 令得到标号的结点属于S，其余结点属于$\bar{S}$
  - 此时$(S, \bar{S})=\{(s,a),(c,b),(c,d)\}$, $C(S, \bar{S})=5$，满足定理5.5.2



Ford-Fulkerson 最大流标号法

- Ford-Fulkerson算法的问题
  - 在算法中，对结点的标号顺序是任意的
  - 即算法可以任选一条s到t的增流路径
  - 每次所选的增流路径并不一定是最好的
  - 算法复杂性可能会依赖于任选的参数
  - 容量是无理数时算法可能失效（需要执行无数步）
- Edmonds-Karp算法
  - 严密的标号算法
  - 每次沿一条最短的增流路径增流



广探法，先标号先检查

第五章 匹配与网络流

- 网络流图
- Ford-Fulkerson最大流标号算法
- 最小费用流
- 互联网中的网络流

最小费用流

- CEO关心的不仅仅是最大流
  - 之前的算法只考虑了最大流，没有考虑网络在运输的时候产生的费用
- 最小费用流问题
  - 如果每条边既有容量，又有单位流量费用，如何从源s以最小费用向目的t发送给定流量w？

最小费用流

- 例5.8.1
  - 一批货物要从工厂运至车站，可以有多条线路进行选择，在不同线路上每吨货的运费不相同，而且每条线路的运货能力有限，怎样运输才能使运费最省？
- 例5.8.2
  - 旅行社安排一批游客，要从甲地飞到乙地，怎样安排才能使旅费最省？
    - 建模：节点代表机场，边表示各个机场间的航班，容量是航班的有效座位数，费用则是机票费

最小费用流

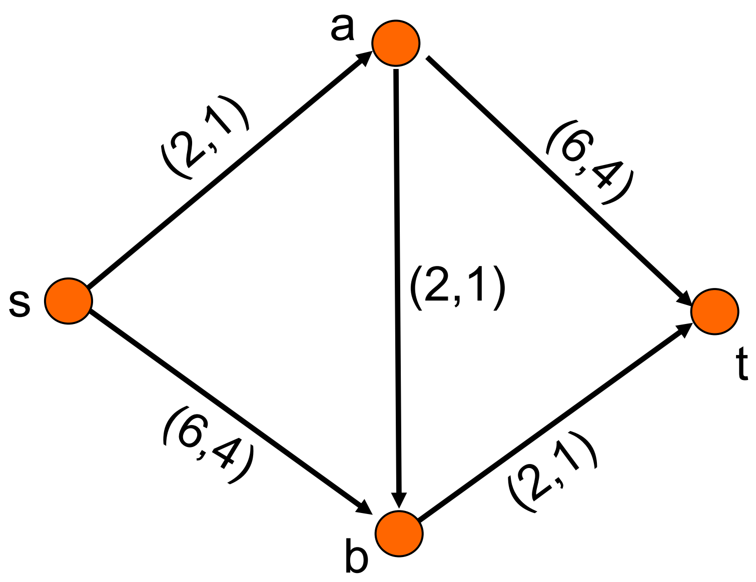
- 设e=(i,j)为网络流图N中的一条边
  - c_{ij}代表该边的容量，a_{ij}表示单位量的费用（即运价），f_{ij}是该边的当前流
- 优化目标和约束条件
  - 最小化费用，给定从s到t的流量w
- 最小费用流问题（建模与代数表示）

$$\min \sum_{eij} a_{ij} f_{ij},\ \text{s.t.}$$

1. $0 \le f_{ij} \le c_{ij}$,
2. $\sum_j f_{ij} = \sum_j f_{ji}, \quad i \ne s,t$
3. $\sum_j f_{sj} = \sum_j f_{jt} = w$

最小费用流

- 解决该问题的一个较好方法是瑕疵算法（需要线性规划知识）
- 简单方案的设计思路
  - 寻找最小费用路径，然后增加流量
  - 优先选取单位流量费用小的边以及道路
  - 把费用看作边的长度
  - 寻找s到t的最短增流路
  - 最终流量达到w，总费用也一般最小



(a_{ij}, c_{ij})：运价和容量

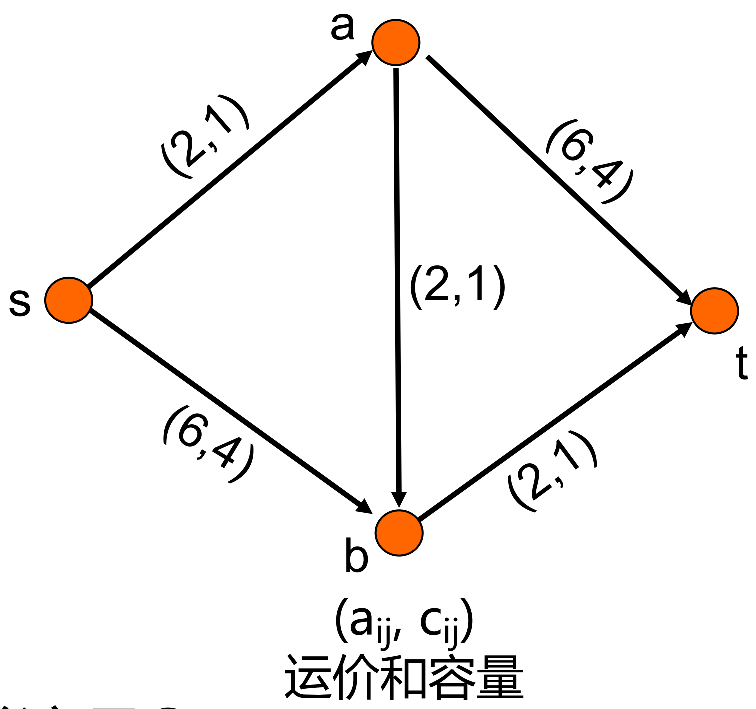
最小费用流

- 算法描述
  - 1. 初始流分布f_0使每条边e都为f(e)=0
  - 2. 当前允许流分布下，修改各边费用：

    $a_{ij}^*=a_{ij},\quad 0\le f_{ij}\le c_{ij}$

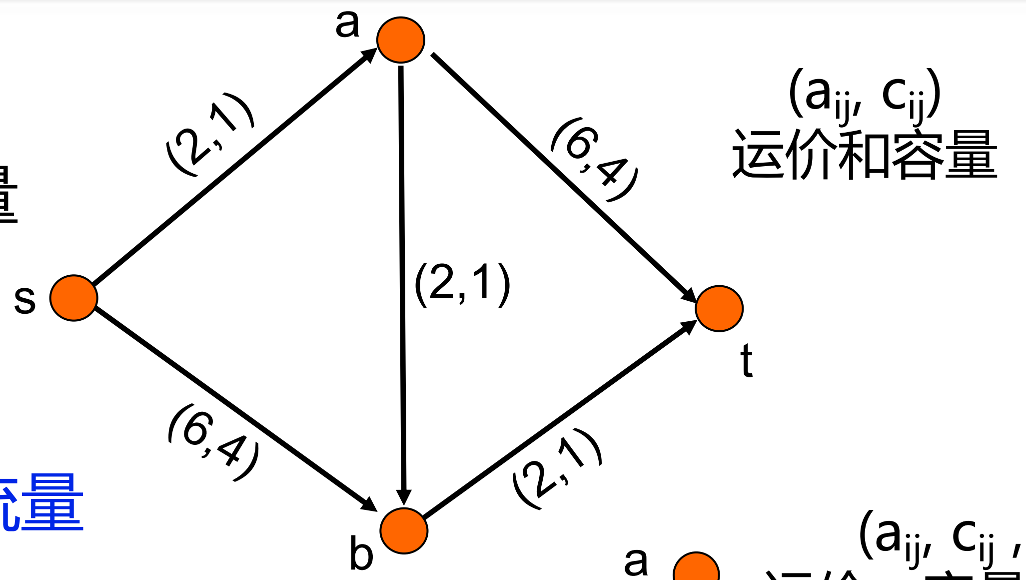
    $a_{ij}^*=\infty,\quad f_{ij}=c_{ij}$　　饱和边

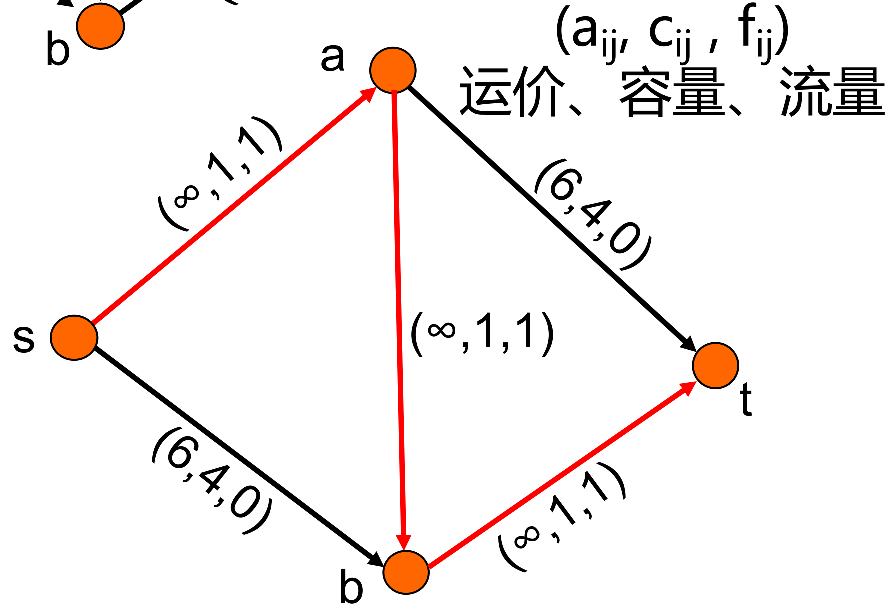
    $a_{ij}^*=-a_{ji},\quad f_{ji}>0$　　鼓励减少使用量
  - 3. 以a_{ij}^*为边长，找s到t的最短增流路，得到增流量δ_t
  - 4. 若$\delta_t+w_0\ge w$,则$\delta_t=w-w_0$，进行最后一次增流，结束；否则转5
  - 5. 进行增流（由δ_t修改允许流），转2



最小费用流

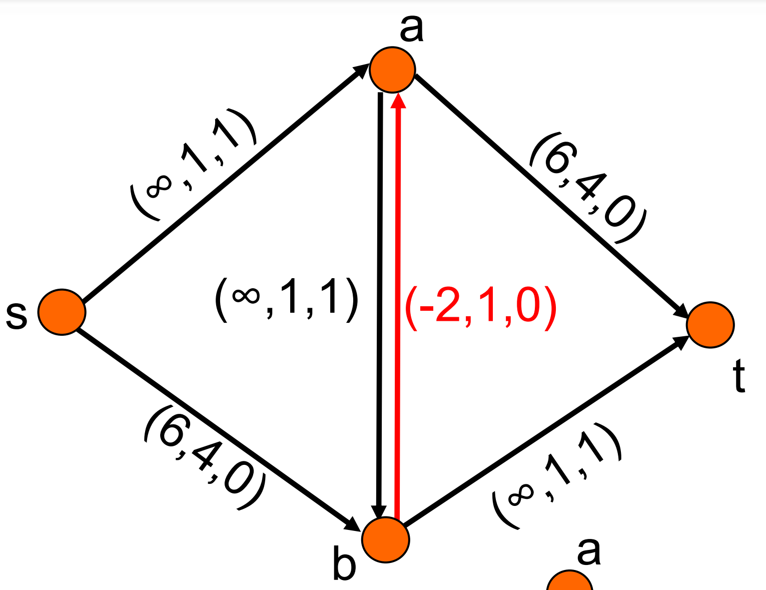
- 例5.8.3
  - 运输网络的每条边都有运价和容量
  - 初始流量$w_0=0$，各边的费用 $a_{ij}^*=a_{ij}$
  - P(s,a,b,t)是当前的最短增流路增流量 $\delta_t=1$，即总流量$w_0=1$
  - 增流后每边第3个权为当前流量
  - 当前总费用为6（对照原图）
  - 修改各边的费用a_{ij}^*

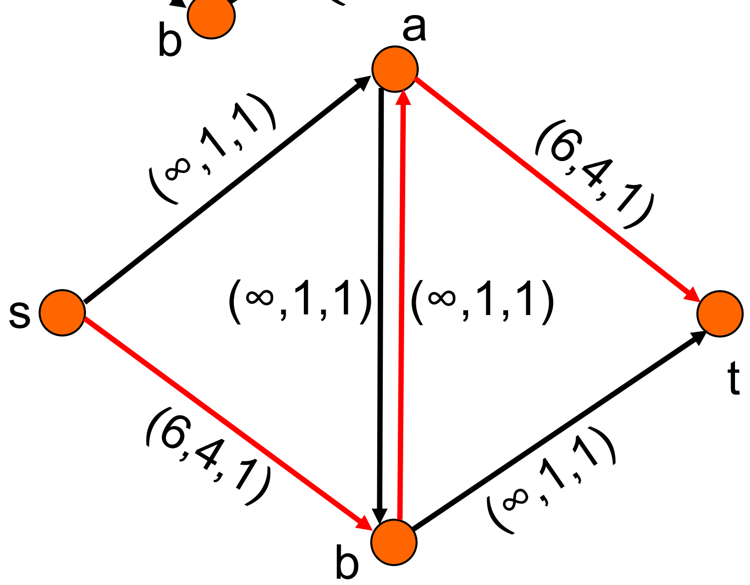




最小费用流

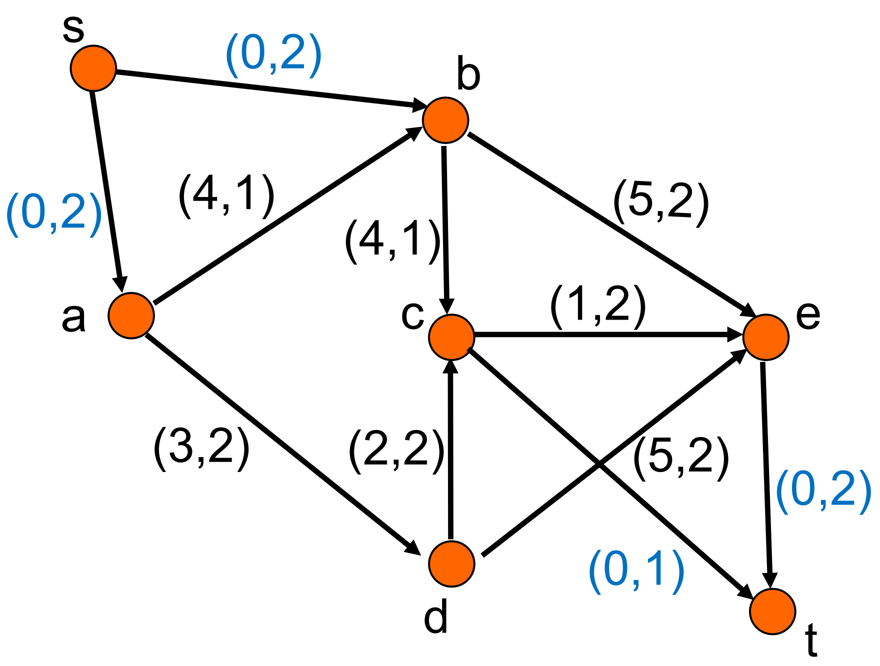
- 例5.8.3（续）
  - 当边(i,j)(i,j≠s,t)的$f_{ij}>0$时，就对应存在一条边(j,i)，并且$a_{ji}^*=-a_{ij}$，$c_{ji}=f_{ij}$，$f_{ji}=0$，再求s到t的最短增流路
  - P=(s,b,a,t)是当前的最短增流路
  - $\delta_t=1$，费用为10
  - 总流量$w_0=2$
  - 总费用$\sum a_{ij}f_{ij}=6+10=16$
  - 原图不存在增流路径，得到最小费用流





最小费用流

- 例5.8.4
  - 多源多汇网络（花费，容量）
    - 发点a,b均可供应2个单位
    - 收点c,e各接收1,2个单位
  - 增设一个超发点s
    - $a_{sa}=0$，$c_{sa}=2$；$a_{sb}=0$，$c_{sb}=2$
  - 增设一个超收点t
    - $a_{ct}=0$，$c_{ct}=1$；$a_{et}=0$，$c_{et}=2$
  - 转为单源单汇

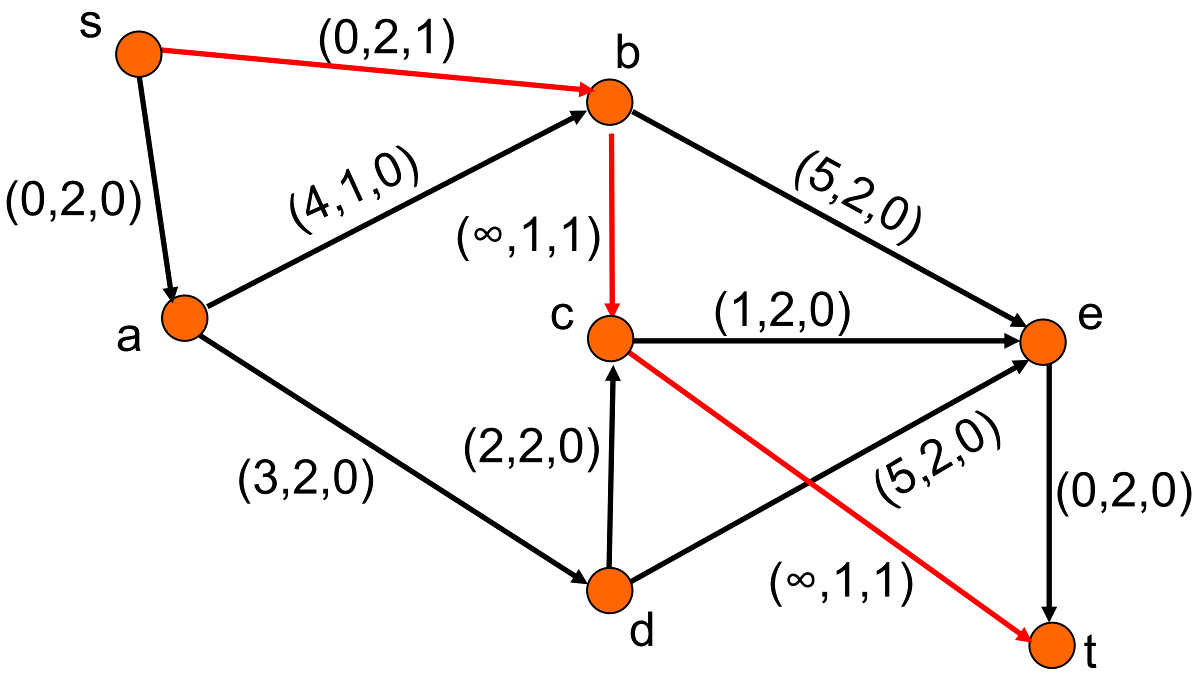


找<s,t>最小花费路径

$w_0=0$，最短增流路径$P_1=(s,b,c,t)$，$\delta_t=1$

最小费用流

- 例5.8.4(续)
  - $w_0=0$，最短增流路径$P_1=(s,b,c,t)$，$\delta_t=1$

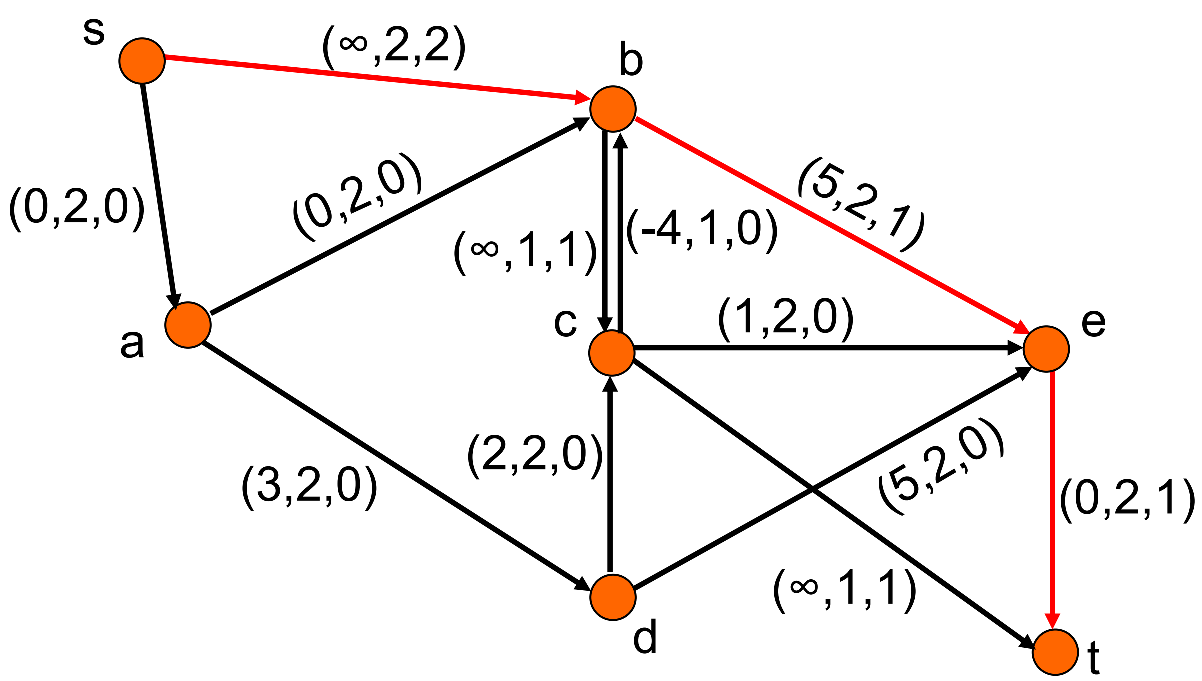


增流修改权值后，再找<s,t>可增流最小花费路径

$w_0=1$，找到最短增流路径$P_2=(s,b,e,t)$，$\delta_t=1$

最小费用流

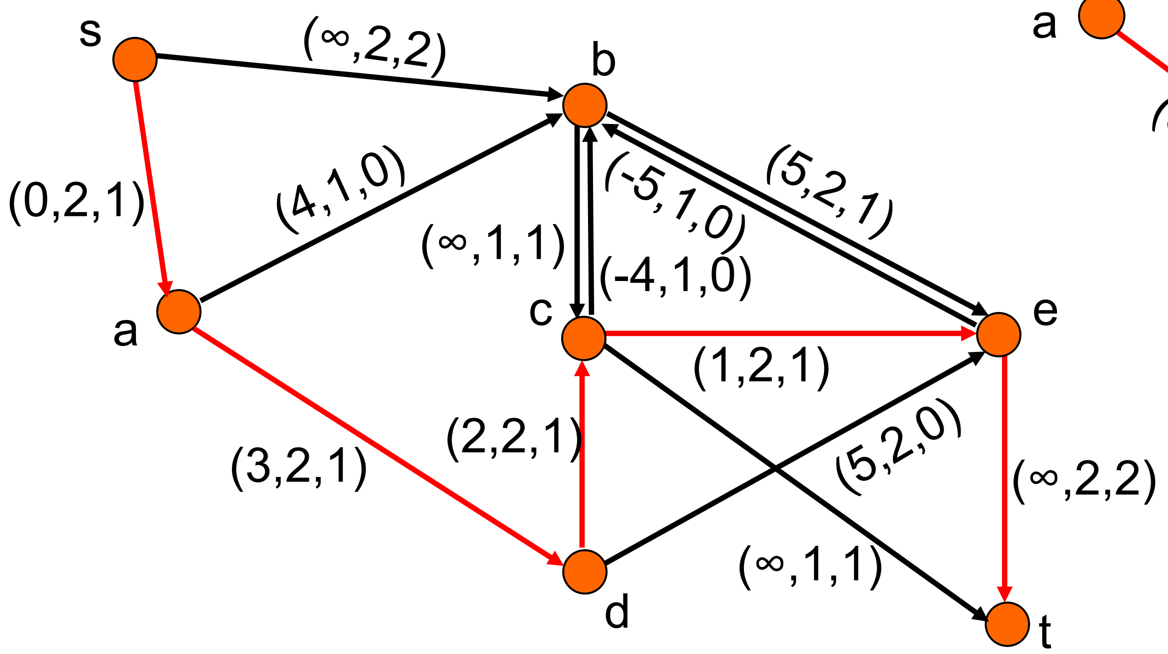
- 例5.8.4(续)
  - 沿(s,b,e,t)增流，再找<s,t>可增流最小花费路径

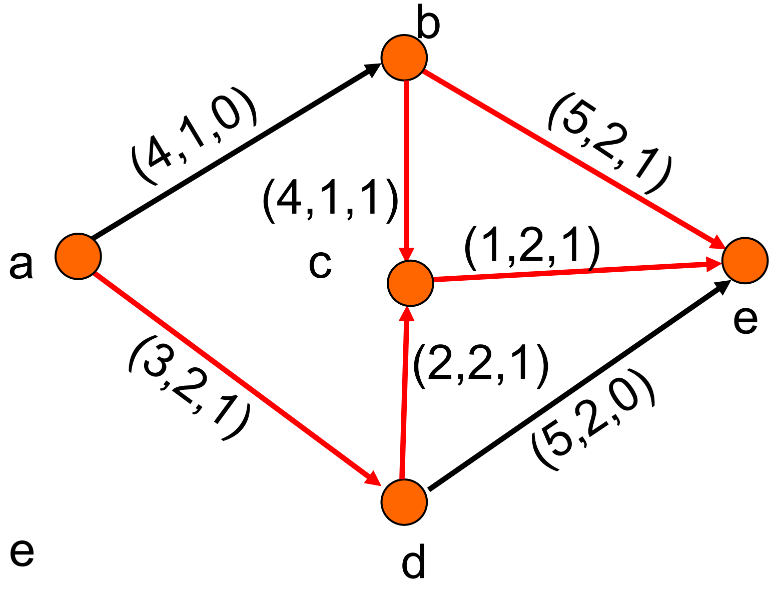


$w_0=2$，最短增流路径$P_3=(s,a,d,c,e,t)$，$\delta_t=1$

最小费用流

- 例5.8.4(续)
  - 沿(s,a,d,c,e,t)进行增流
  - $w_0=3$，完成



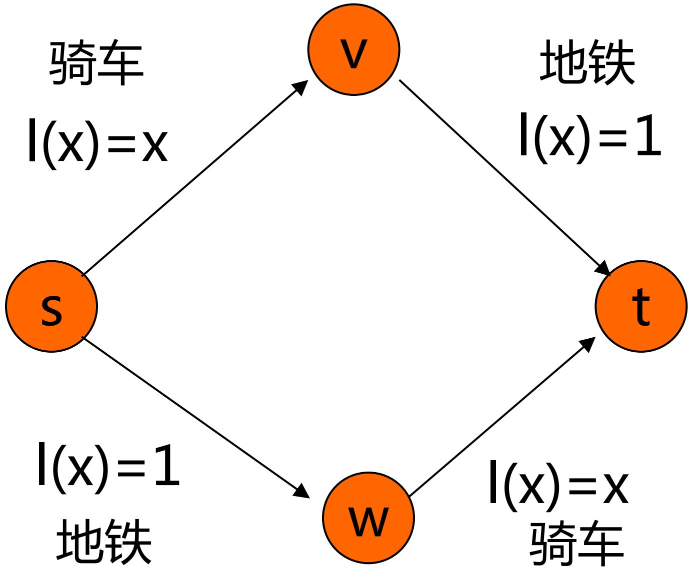


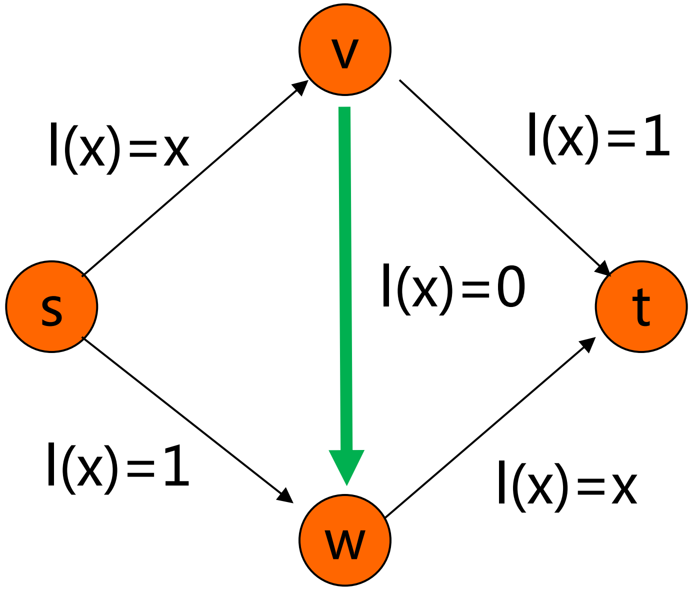
第五章 匹配与网络流

- 网络流图
- Ford-Fulkerson最大流标号算法
- 最小费用流
- 互联网中的网络流

图论中的应用趣题

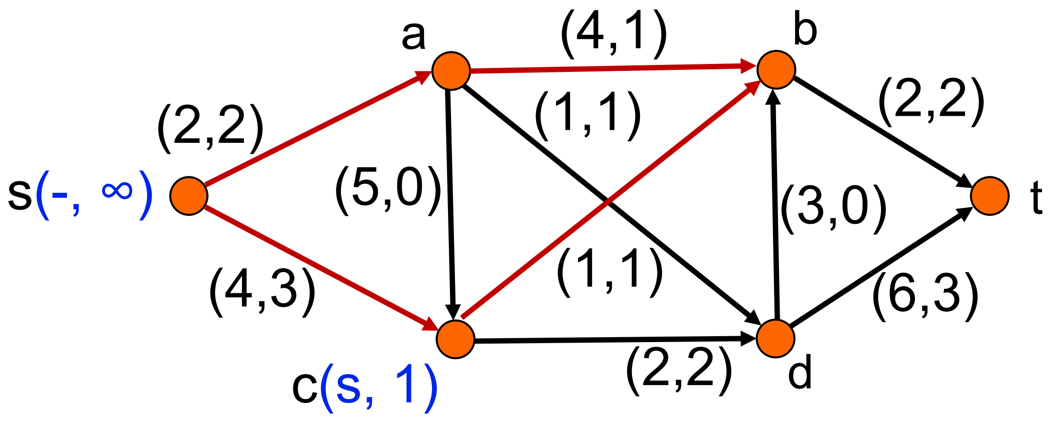
- 城市修路问题：道路通过时延为l(x)，其中x为该边的流量大小





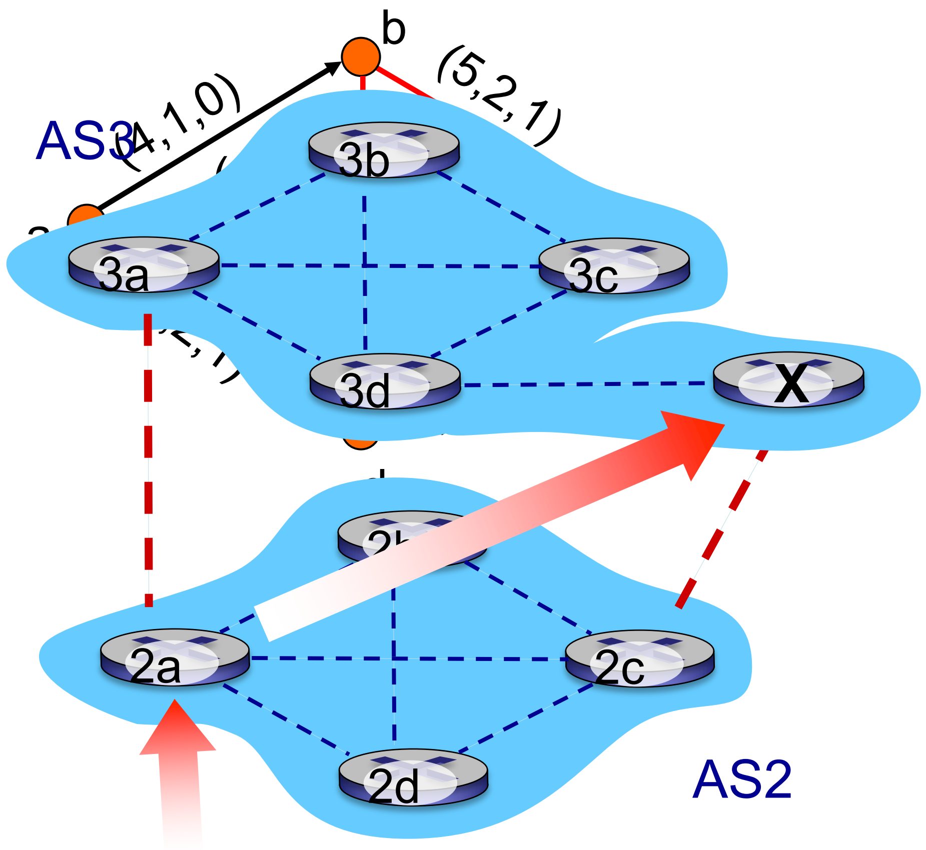
最大流与互联网路由

- 互联网路由
  - 算法 v.s. 协议，集中式 v.s. 分布式
  - 互联网设计原则：面向可扩展的端到端原则
  - 互联网采用可扩展的分布式路由（甚至是单一最佳路由）



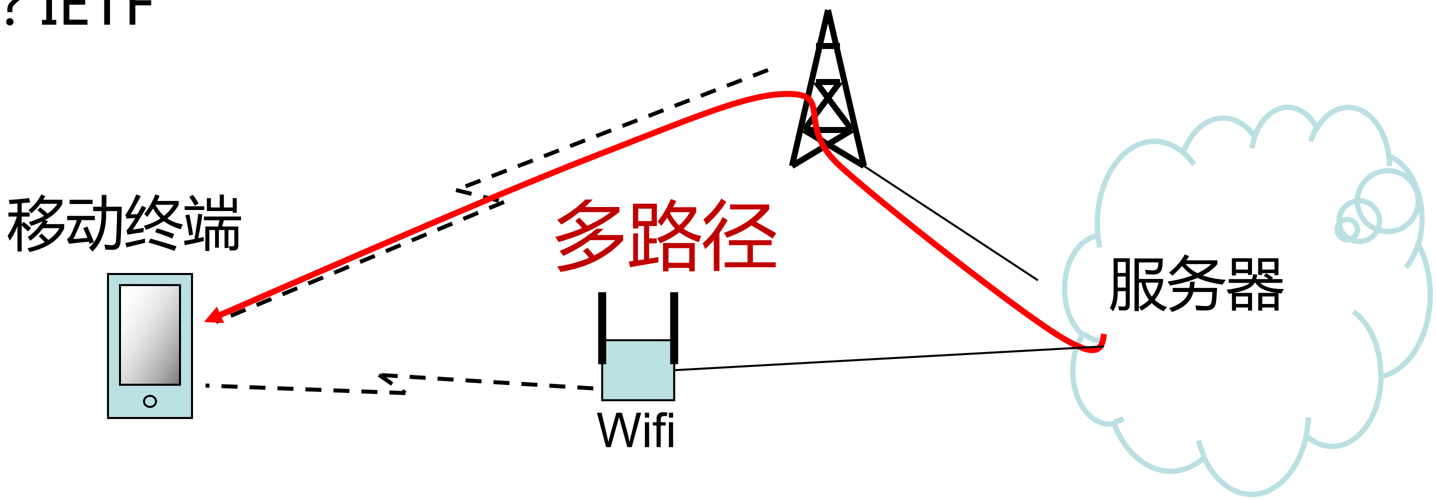
互联网路由优化

- 互联网路由中设置的“费用”
  - 路由信息协议RIP
    - 只关心跳数——最小跳数路由
  - 开放式最短路径优先协议OSPF
    - 通过设置链路权值w，计算最短路
    - 权值设置为链路带宽的倒数
    - 多条“最短路”之间有可能同时使用
  - 域间路由协议BGP
    - 权值设置考虑“策略”（“热土豆”）



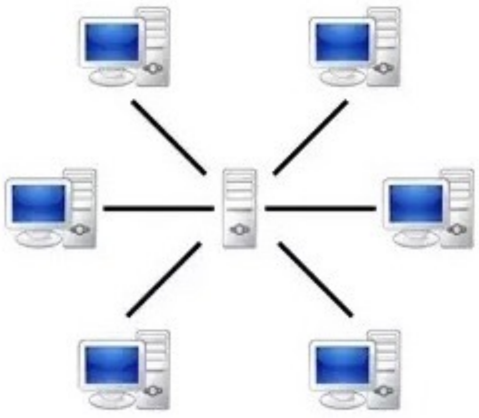
传输协议：从单径到多径

- 互联网的核心承载协议
  - 负责端到端的传输控制，单路径传输
  - 如何努力提高流量？拥塞控制是核心
- 多路径传输
  - 多路径之间的调度，避免浪费好路径
  - 如何部署？IETF

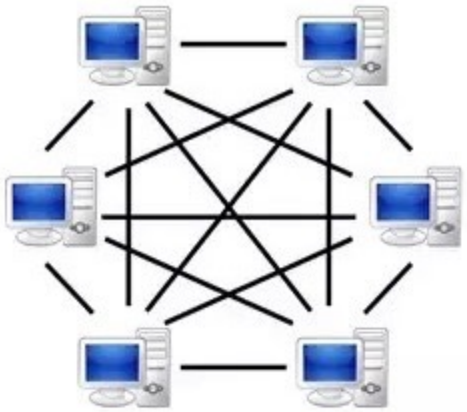


P2P网络

- C/S模式：client-server
  - 从中心服务器下载，可扩展性？
  - 不赚钱就要省钱：爱奇艺的大红包
- P2P网络（应用层网络）
  - 没有中心服务器，同时从多个机器下载/上传
  - 谁有什么内容？我从哪里取最好？



Server-based



P2P-network

迅雷路由器
全球第一台
会赚钱的路由器

总结：网络流

- 网络流图概念
  - 掌握概念，学习玩命简化的建模，分析网络流图的增流思路
  - 掌握割切的概念和最大流定理
- Ford-Fulkerson最大流标号算法
  - 通过寻找从s到t的增流路径，不断增流
  - 标号的设计思路：增流路径、流量大小、前后方向
- 最小费用流
  - 掌握概念，了解基本求解思路
- 了解互联网中的网络流

本周作业

- P137习题六
  - 最大流和最小割切：第2题
  - 网络流图的证明题：第4题
  - 补充题1：假如我是欧拉
    - 中午12:15六教下课，千军万马去清芬园吃饭
    - 请尝试分别从单个学生、班级整体和学校的角度，给出不同的优化目标，并列举可能需要考虑的多种约束条件
  - 补充题2：假如我是欧拉
    - 总结创新或著书立说的基本方法（即本书中有哪些一步步创新的思路），并给出具体示例。不少于3条基本方法。

长期作业

- 第13周周一课前提交：图论应用技术报告PPT
  - 给出实际问题，进行图论建模，尝试设计求解思路并分析优缺点
  - 8页以内PPT，不用编程序，不用细节推导
- 制作PPT小技巧
  - 技术思路，难点，思路清晰，重点明确（不要细节）
  - 活用例子，图文并茂，内容简洁（避免主谓宾）
  - 尽量使用PPT模板（活用Tab键）
  - 站在听众角度，引起思考和互动
  - 常见问题：文字太多、字体太小、颜色不清、布局太乱

文字换行了吗？

- 本学期最后一次课前：完成图论内容的初步复习