# 影响力最大化1：degree discount

覃雄派

# 影响力最大化1：degree discount

- 首次提出influence maximization

Maximizing the Spread of Influence through a Social Network

David Kempe, Jon Kleinberg, Eva Tardos

Cornell University

KDD 2003

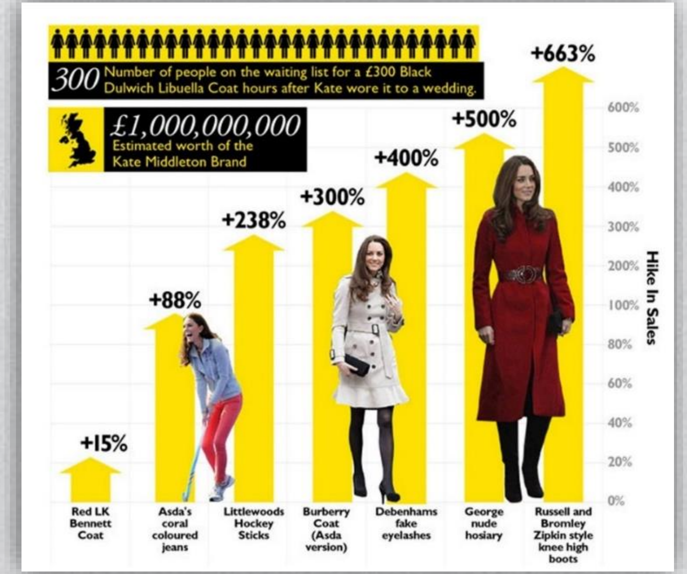
影响力最大化1：
degree discount

# 提纲

- 社交影响力
- IC传播模型
- 影响力最大化及其困难
- 贪心算法

# 影响力最大化1：degree discount

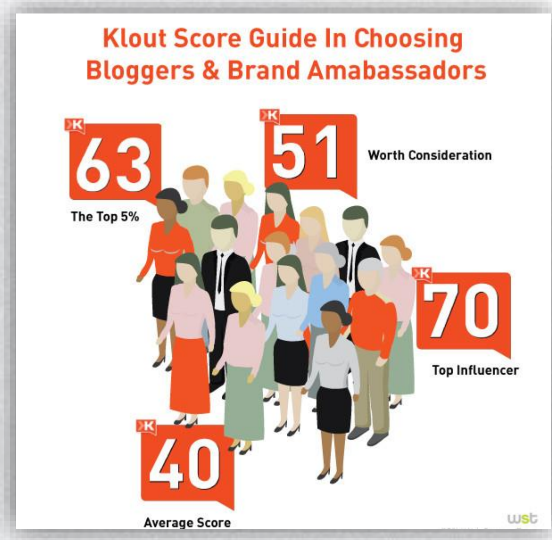
- 社交影响力（Social Influence）
  - 凯特王妃效应：穿衣品味影响力拉动百亿时装消费
  - 最佳的广告推广形式来自熟人之间的口耳相传




Recommendations from people known (90%)

# 影响力最大化1：degree discount

- 社交影响力（Social Influence）
  - 在线社交网络迅猛发展
  - 在线社交影响力：度量用户在社交网络上对其听众有多大影响

# 影响力最大化1：degree discount

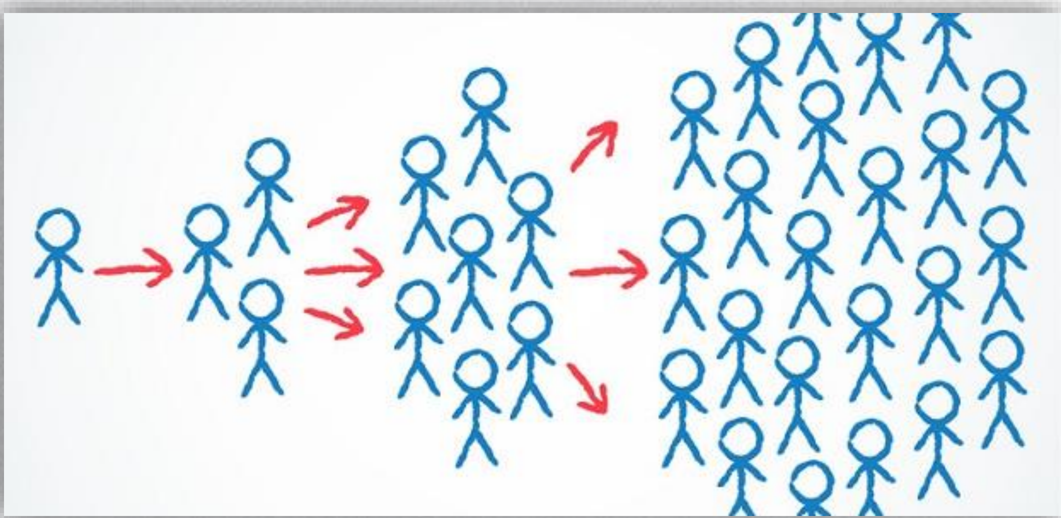
- 社交影响力（Social Influence）：应用
  - Social influence occurs when a person's emotions, opinions, or behaviors are affected by others

-- From 



主要应用

病毒式营销



# 影响力最大化1：degree discount

- 社交影响力（Social Influence）：应用
  - 谣言与虚假信息防控

- 世界经济论坛十大全球危机（2014）
  - 1. Rising societal tensions in the Middle East and North Africa
  - 2. Widening income disparities
  - 3. Persistent structural unemployment
  - ...
  - *10. The rapid spread of misinformation online*

# 影响力最大化1：degree discount

- 真实应用实例：微信朋友圈广告
  - 微信朋友圈广告是典型的feeds流广告
  - feeds广告就是与内容混排在一起的广告
  - 最不像广告的广告，长得最像内容的广告
  - Feeds广告操作性简单，打扰性低，已经成为移动互联网时代主流的广告形式
- 建立在用户行为记录和大数据分析基础上
  - 个性化推荐
- 微信朋友圈广告中的社交影响力
  - 微信朋友圈广告具有互动属性

# 影响力最大化1：degree discount

- 真实应用实例：微信朋友圈广告中的社交影响力
  - 微信朋友圈广告具有互动属性

广告互动
- 点赞
- 评论
- @好友

节点价值 = × 点赞概率 + × 评论概率 + × 个人影响力

社交关键节点先投放，逐层影响其它节点

# 影响力最大化1：degree discount

- 信息传播与影响力最大化：应用场景
  - 病毒营销
  - 互联网不实信息的控制
  - 政治宣传
  - 等……

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 节点度数的分布：在大多数真实图（如社交图）中，节点度数呈现什么分布?
  - Power Law幂律分布，如图所示
    - 度量分布函数$P$（$K$），其中$K$为节点度数， $P$（$K$）表示度数为$K$的节点出现的频率
    - Power Law度数分布的直观解释 – 长尾分布：大多数节点的度数都很低，只有少部分节点的度数很高，即对他人能够产生较大影响力
    - 著名的“二八定律”

# 影响力最大化1：degree discount

- 案例分析：公众号的幂律分布
  - 大号：既得利益者，上有远虑
  - 腰部账号：正值当年，尚能饭否
  - 尾部账号：生无可恋，回天乏术

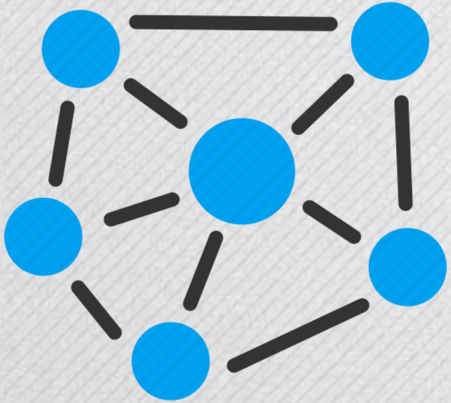
| 类型 | 定义 | 所占百分比 | 现状概括 | 重点事项 |
|---|---|---|---|---|
| 头部 | >100万粉丝 | 1.80% | 变现 | 帮别人卖：接广告<br>自己卖：实物类，虚拟类（知识付费，会员） |
| 腰部 | 1-100万 | 19.60% | 折腾 | 涨粉+变现 |
| 尾部 | <1万 | 78.60% | 挣扎 | 内容+涨粉 |


http://www.niaogebiji.com/article-16711-1.html

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- IC传播模型
  - 如何建模社交网络信息传播
  - 社交网络信息传播模型



在此讲述
Independent Cascades (IC) Model

❖ Compartmental models (SI, SIR, SIS, etc.)
[Kermack and McKendrick, 1927]

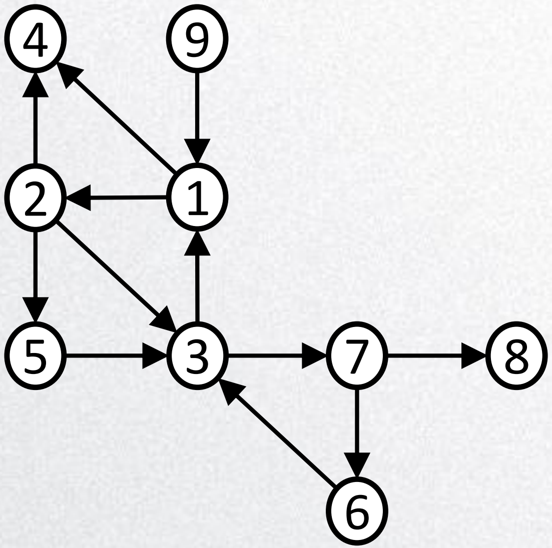
❖ Rumor spreading models (DK, MT)
[Daley and Kendall 1964, Maki 1973]

❖ Independent Cascades Model
[Kempe et al., 2005]

❖ Threshold Model, Complex Contagion
[Granovetter 1979, Centola 2010]

# 影响力最大化1：degree discount

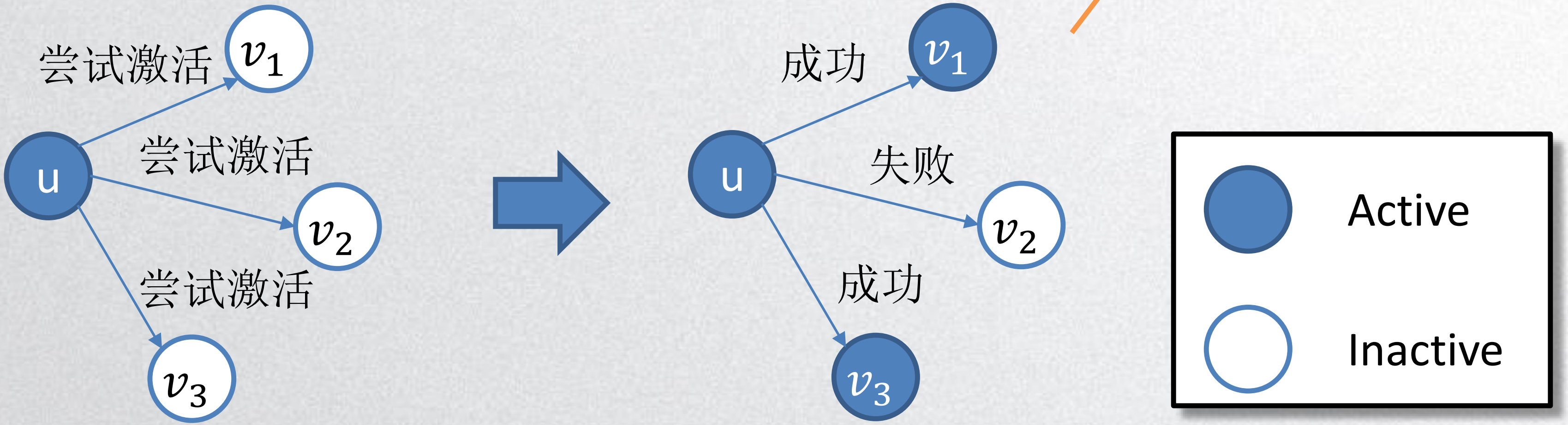
- IC传播模型
  - 建立有向图模型
  - 使用有向图对社交网络上的信息传播进行建模
  - 社交网络：有向图$G = (V, E)$
    - 点集$V$：表示社交网络用户的集合
    - 边集$E$：表示社交网络上的信息传播关系
  - 例子：
    - 如果A在微博上关注了B
    - 那么存在一条有向边从B指向A

# 影响力最大化1：degree discount

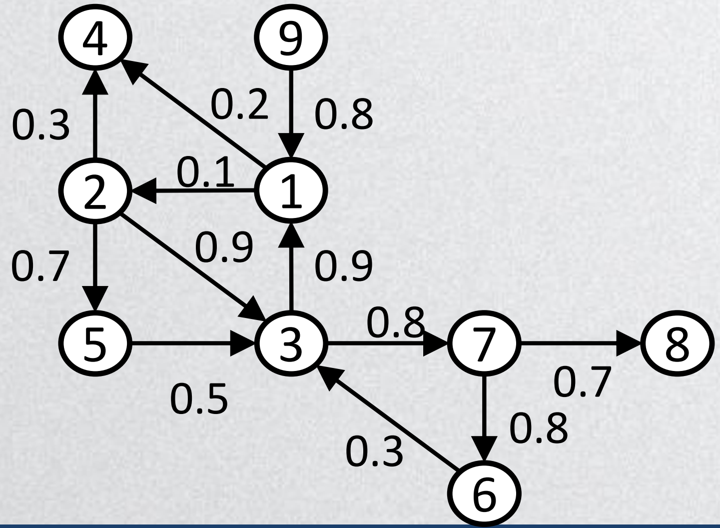
- Independent Cascades模型(IC模型)
- IC模型考虑图上的节点有两个状态
  - Active: 即成功地被影响到
  - Inactive: 即当前还未被影响到
- IC模型考虑每一个节点有一次机会去激活它的邻居

引入概率度量成功/失败的可能性



# 影响力最大化1：degree discount

- Independent Cascades模型
- 假设用户u到v之间有一个概率值，称为影响概率
  - 概率值越大，v相信u传播消息的可能性就越大
  - 概率值可以通过u和v的互动历史学习得到



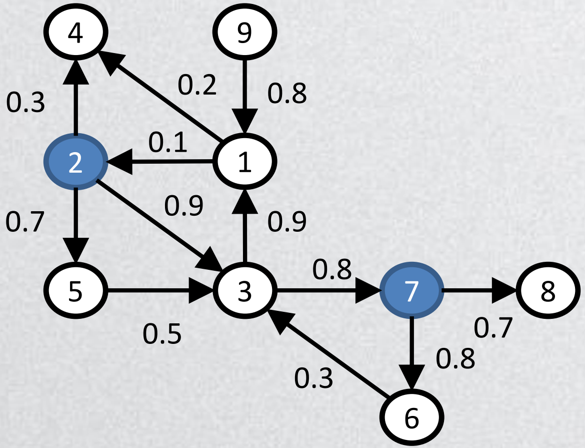

影响概率
Influence Probabilities

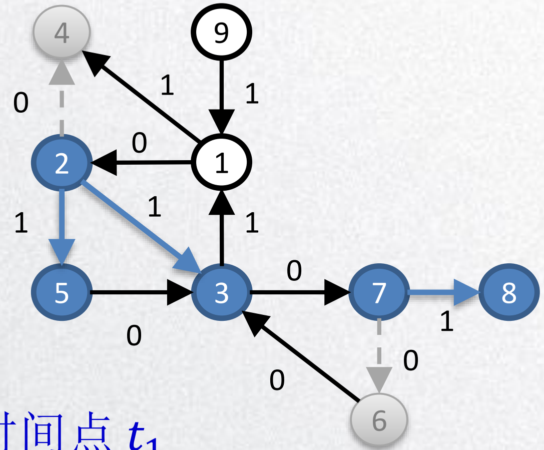
# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 影响力最大化及其困难
  - 找到种子节点集合，使得其影响范围(Influence spread)最大
  - 假设选择$u_2$和$u_7$作为初始的传播者，称为种子
  - 种子用户**按照影响概率**去“激活”其邻居

有的邻居被激活
有的邻居没有被激活

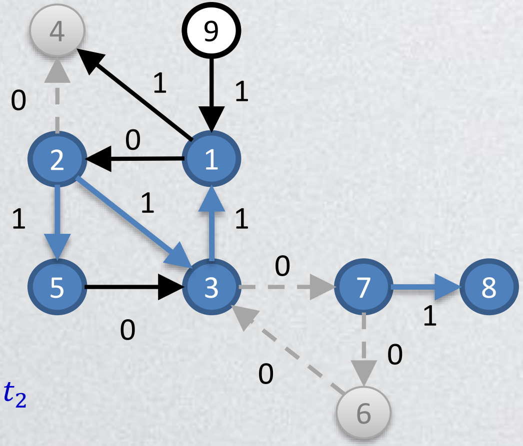



时间点 $t_1$

# 影响力最大化1：degree discount

- 影响力最大化及其困难
  - 找到种子节点集合，使得其影响范围(Influence spread)最大
  - 被激活的邻居继续影响他们的好友
    - 注：每个用户只有一次机会去激活其好友
  - 直到没有新的用户被激活，停止




时间点 $t_2$

✓ $u_1$
✓ $u_2$
✓ $u_3$
✓ $u_5$
✓ $u_7$
✓ $u_8$

得到$u_2$和 $u_7$的影响范围

上述过程重复多次，求出平均的影响范围

# 影响力最大化1：degree discount

- 影响力最大化Influence Maximization及其困难
  - 找到信息传播中的“关键种子节点”

选择最优种子集合$S^*$

使得其期望的影响范围最大化

$$S^* = \arg_S \max \sigma(S)$$

一般是有预算的，不能无限扩大集合$S^*$

影响力范围$\sigma(S)$：度量以$S$为种子，使用IC模型，最后平均激活多少节点

# 影响力最大化1：degree discount

- 影响力最大化Influence Maximization及其困难
  - 找到信息传播中的“关键种子节点”

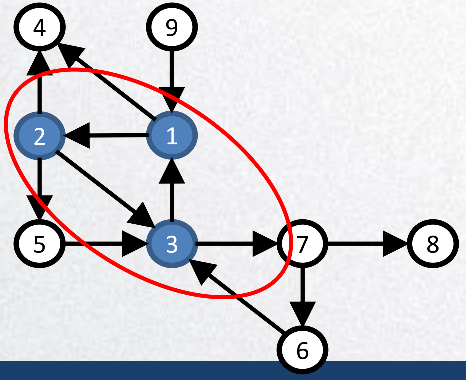
选择最优种子集合$S^*$

使得其期望的影响范围最大化

$$S^* = \arg_S \max \sigma(S)$$

一般是有预算的，不能无限扩大集合$S^*$

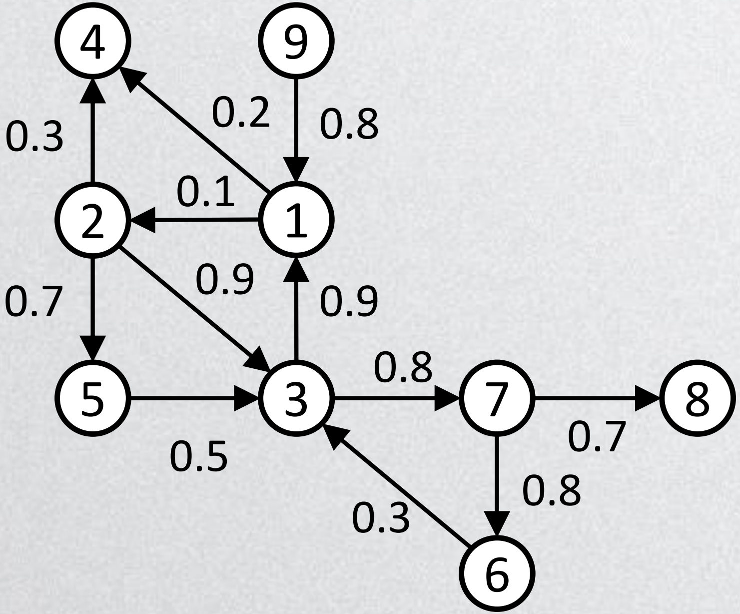
比如限定节点数量为3，得到最优种子节点集合为{1,2,3}




影响力范围$\sigma(S)$：度量以$S$为种子，使用IC模型，最后平均激活多少节点

# 影响力最大化1：degree discount

- 影响力最大化及其困难
  - 找到种子节点集合，使得其影响范围(Influence spread)最大




如果只让你选择2个种子节点，**你会怎么选?**

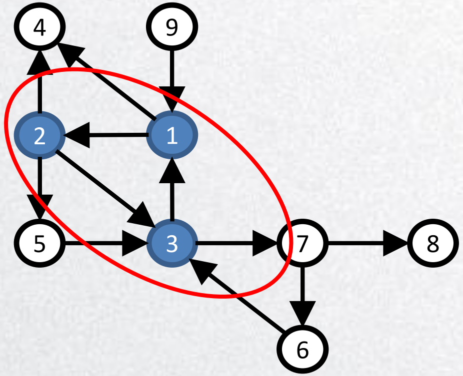
1.对所有的2个结点的组合，使用前述办法计算其影响范围

从这2个节点开始激活其他节点，直到没有节点被激活为止，**多次做，求平均(代价很大)**

2.从这些组合中，找出影响范围最大的

# 影响力最大化1：degree discount

- 影响力最大化及其困难
  - 找到种子节点集合，使得其影响范围(Influence spread)最大
  - 即找到信息传播中的“关键种子节点”
  - 这个问题有多复杂?
    - NP难问题：不存在一个时间复杂度与节点个数呈多项式关系的算法
  - 计算任意$k$个点$S$的影响力$\sigma(S)$
    - 需要模拟多次之前的IC模型
    - 时间复杂度高（#P难问题）



# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 贪心算法：基本想法
  - 使用贪心的策略选择种子节点
  - 每次选择使当前$\sigma(S)$增益最大的节点

1.Initialize $S$=∅
2.For $i$=1,......,$k$ do:
 1) Select the node $u \leftarrow \arg\max_{w \in V-S}(\sigma(S \cup \{w\}) - \sigma(S))$;
 2) $S \leftarrow S \cup \{u\}$
3.Return $S$.

性质：如果$\sigma(S)$满足单调且次模，上述贪心算法有$1-\frac{1}{e}$的近似比

# 影响力最大化1：degree discount

- 贪心算法
  - 每次迭代，都尝试增加一个节点
  - 假设已有的集合为S
    - 增加节点为w
    - 从这$S \cup \{w\}$的节点出发，开始激活其他节点，直到没有节点被激活为止，多次做求平均$\sigma(S \cup \{w\})$
    - $\sigma(S \cup \{w\})$减去已有的$\sigma(S)$
  - 看看那个w导致的增量最大，就选它

代价很大

# 影响力最大化1：degree discount

- 贪心算法：基本想法
  - 使用贪心的策略选择种子节点
  - 每次选择使当前$\sigma(S)$增益最大的节点

1.Initialize $S$=∅
2.For $i$=1,……,$k$ do:
  1) Select the node $u \leftarrow \mathrm{arg}max_{w \in V-S}\left(\boxed{\sigma(S \cup \{w\}) - \sigma(S)}\right)$;
  2) $S$←$S$∪{$u$}
3.Return $S$.

如何高效地计算$\sigma(S)$增益

# 影响力最大化1：degree discount

- 解决思路：启发式算法
  - 启发式算法是相对于最优化算法提出的
    - 最优算法求得该问题每个实例的最优解
  - 一个基于直观或经验构造的算法
    - 在可接受的花费（指计算时间和空间）下给出待解决优化问题的一个可行解
    - 该可行解与最优解的偏离程度一般不能被预计

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

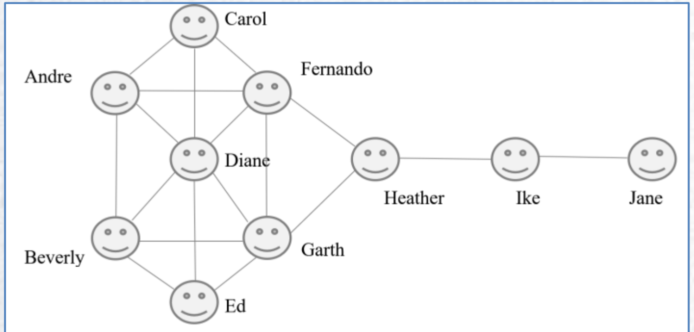
- 贪心算法
  - 效率不高
  - 需要引入一些heuristics(启发信息、启发规则)

- 可能的启发式算法
  - 选择Degree Centrality最大的$k$个节点
  - 选择PageRank分数最大的$k$个节点

# 影响力最大化1：degree discount

- 贪心算法
  - 效率不高
  - 需要引入一些heuristics(启发信息、启发规则)

- 可能的启发式算法
  - 选择Degree Centrality最大的$k$个节点
    - 有问题
  - 选择PageRank分数最大的$k$个节点



Kite 网络(Kite Network)

在这个网络中，Diane的度中心性最高；Fernando的closeness中心性最高，到其他节点的距离最短；Heather 的Betweenness中心性最高，连接这个网络的2个部分；那个节点的影响力最大呢？更加有利于影响力的传播呢？

# 影响力最大化1：degree discount

- 贪心算法
  - 效率不高
  - 需要引入一些heuristics(启发信息、启发规则)

- 可能的启发式算法
  - 选择Degree Centrality最大的$k$个节点
    - 有问题
  - 选择PageRank分数最大的$k$个节点
    - 有问题

> - page rank高的节点，表示大家都给你投票，和你的传播范围广不一定是一回事；
> - 入口网站有时候比权威内容网站，传播效果要好

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 在大多数真实图（如社交图）中，节点度数呈现什么分布?
  - Power Law度数分布：度量分布函数$P$（$K$），其中$K$为节点度数，$P$（$K$）表示度数为$K$的节点出现的频率

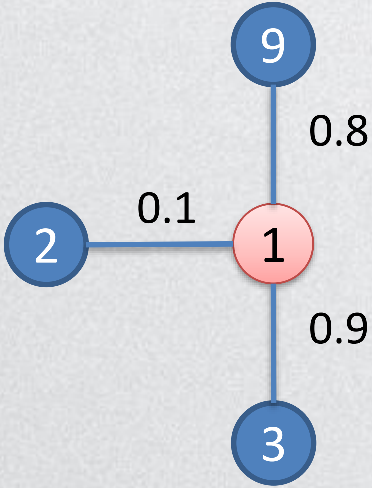
  - Power Law度数分布的直观解释
    - 长尾分布：大多数节点的度数都很低，只有少部分节点的度数很高，即对他人能够产生较大影响力

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 基于节点度数的解决思路
- 基本想法
  - 节点只对其直接邻居产生影响
  - 选择的种子节点的影响范围避免重叠

为了表达简便，我们考虑无向图，即两点$u$和$v$影响对方的概率相同




分类讨论：

- 1没有activated
- 1已经activated

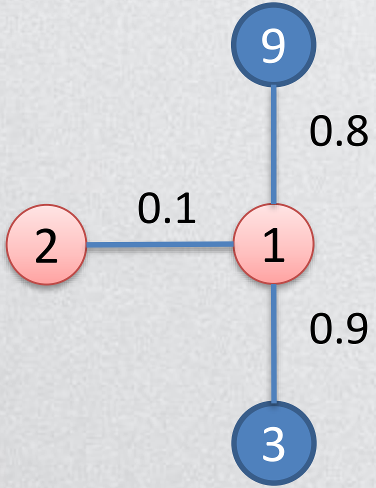
$$\Delta S = 1 + d_v p$$

选择1，$\Delta\sigma(S) = 1 + (1\ 1\ 1)\begin{pmatrix}0.8\\0.9\\0.1\end{pmatrix} = 1 + .8 + .9 + .1 = 2.8$

# 影响力最大化1：degree discount

- 基于节点度数的解决思路
- 基本想法
  - 节点只对其直接邻居产生影响
  - 选择的种子节点的影响范围避免重叠

为了表达简便，我们考虑无向图，即两点$u$和$v$影响对方的概率相同




分类讨论：
- 1没有activated
- 1已经activated

已知$S=\{2\}$，再选择1，$\Delta\sigma(S)=?$

节点1的影响，因为2的已经引入，打了折扣！

$$\Delta S=\boxed{(1-p)^{t_v}}\cdot(1+(d_v-t_v)\cdot p)$$

$$\begin{aligned}\Delta\sigma(S)&=(1-0.1)^1*\ \ (1+(1\ 1)\begin{pmatrix}0.8\\0.9\end{pmatrix}\\&=\ \ (1-0.1)\ \ (1+0.8+0.9)\end{aligned}$$

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

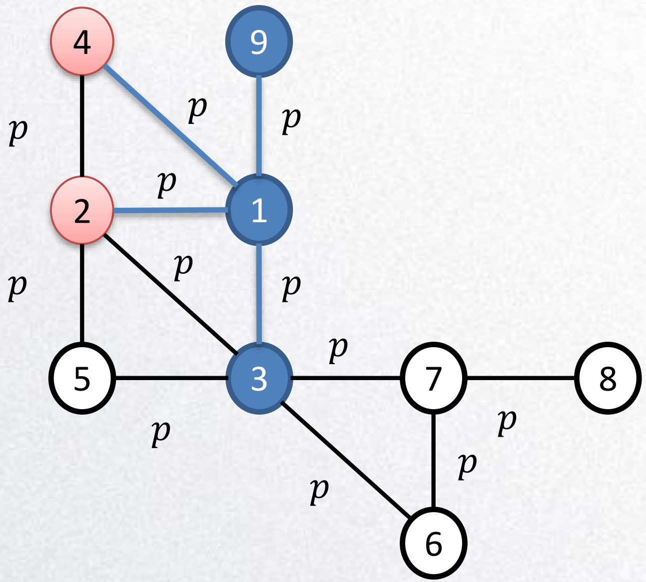
- Degree Discount算法
- 基本想法
  - 出度越高的点，影响力越大
  - 选择的种子节点的影响范围避免重叠

1. Initialize $S = \emptyset$
2. Compute the out-degree of each node;
3. For $i = 1, \ldots \ldots, k$ do:
   1) Select the node $u$ with highest remaining degrees into $S$;
   2) For each vertex $v \in V \setminus S$ such that $v$ is a neighbor of $u$ do:
      a) Discount the degree of $v$ according to edge probability $p_{u,v}$;
4. Return $S$.

如何Discount?

# 影响力最大化1：degree discount

- Degree Discount算法
- 对问题做简化：
  - 每条边的影响概率都为$p$
  - 定义$d_v$是节点$v$的度数，即邻居节点的个数
  - 定义$t_v$是v的邻居中已经选择进入种子集合的个数
  - 定义$dd_v$为折扣后的度数，初始时设定$dd_v = d_v$



# 影响力最大化1：degree discount

- Degree Discount算法
- 给定当前的种子集合$S$，如果将节点$v$加入到$S$中，$\Delta S = ?$
  - 如果$t_v = 0$，$\Delta S = 1 + d_v p$
  - 如果$t_v \neq 0$，可以推出

THEOREM *In the IC model with propagation probability $p$, suppose that $d_v = O(1/p)$ and $t_v = o(1/p)$ for a vertex $v$. The expected number of additional vertices in $Star(v)$ influenced by selecting $v$ into the seed set is:*

$$\boxed{1 + (d_v - 2t_v - (d_v - t_v)t_v p + o(t_v)) \cdot p}$$

**Efficient Influence Maximization in Social Networks**

Wei Chen
Microsoft Research Asia
Beijing, China
weic@microsoft.com

Yajun Wang
Microsoft Research Asia
Beijing, China
yajunw@microsoft.com

Siyu Yang
Dept. of Computer Science
Tsinghua University
Beijing, China
siyu.yang@gmail.com

这个公式是怎么来的？
有什么假设？
请见后文

https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/weic-kdd09_influence.pdf

# 影响力最大化1：degree discount

- 证明

THEOREM *In the IC model with propagation probability $p$, suppose that $d_v = O(1/p)$ and $t_v = o(1/p)$ for a vertex $v$. The expected number of additional vertices in $Star(v)$ influenced by selecting $v$ into the seed set is:*

$$1 + (d_v - 2t_v - (d_v - t_v)t_v p + o(t_v)) \cdot p$$

PROOF. Let $S_v$ be the set of $t_v$ neighbors of $v$ that have been selected as seeds. The probability that $v$ is influenced by its immediate neighbors is $1 - (1-p)^{t_v}$. In this case, selecting $v$ as a seed does not contribute additional influence in the graph.

If $v$ is not influenced by any of the already selected seeds, which occurs with probability $(1-p)^{t_v}$, then the additional vertices in $Star(v)$ influenced by selecting $v$ into the seed set include: (a) $v$ itself with probability 1; and (b) each $u$ in the remaining $d_v - t_v$ neighbors with probability $p$. Thus the additional vertices in $Star(v)$ influenced by $v$ is $1 + (d_v - t_v) \cdot p$. Hence the overall expected number of additional vertices in $Star(v)$ influenced by $v$ is

$$\begin{aligned}
&(1-p)^{t_v} \cdot (1 + (d_v - t_v) \cdot p) \\
&= (1 - t_v p + o(t_v p)) \cdot (1 + (d_v - t_v) \cdot p) \\
&\qquad\qquad \{\text{since } t_v p = o(1)\} \\
&= 1 + (d_v - 2t_v)p - (d_v - t_v)t_v p^2 + o(t_v p) \\
&\qquad\qquad \{\text{since } (d_v - t_v)p = O(d_v p) = O(1)\} \\
&= 1 + (d_v - 2t_v - (d_v - t_v)t_v p + o(t_v))p.
\end{aligned}$$

□

1.注意：理解这段证明的关键是p很小
2.(1-p)的某次方做展开以后得到加上$1-t_v p$以及一个很小的数

比如,p=0.01
$(1-p)^2 = 1-2p + p^2$
最后那1项很小

$(1-p)^3 = 1-3p+3p^2-p^3$
后面那2项都很小

# 影响力最大化1：degree discount

- Degree discount算法

**Algorithm** DegreeDiscountIC$(G, k)$

```
initialize S = ∅
for each vertex v do
    compute its degree d_v
    dd_v = d_v
    initialize t_v to 0
end for
for i = 1 to k do
    select u = arg max_v {dd_v | v ∈ V \ S}
    S = S ∪ {u}
    for each neighbor v of u and v ∈ V \ S do
        t_v = t_v + 1
        dd_v = d_v − 2t_v − (d_v − t_v)t_v p
    end for
end for
output S
```

$d_v$较大，$t_v$较小

For example, for a node v with $d_v = 200$, $t_v = 1$, and $p = 0.01$ (parameters similar to our experimental graphs), we should discount v's degree to about 196：
$d_v - 2t_v - (d_v - t_v)t_v p = 200 - 2*1 - (200-1)*1*0.01 = 200 - 2 - 1.99 = 196$

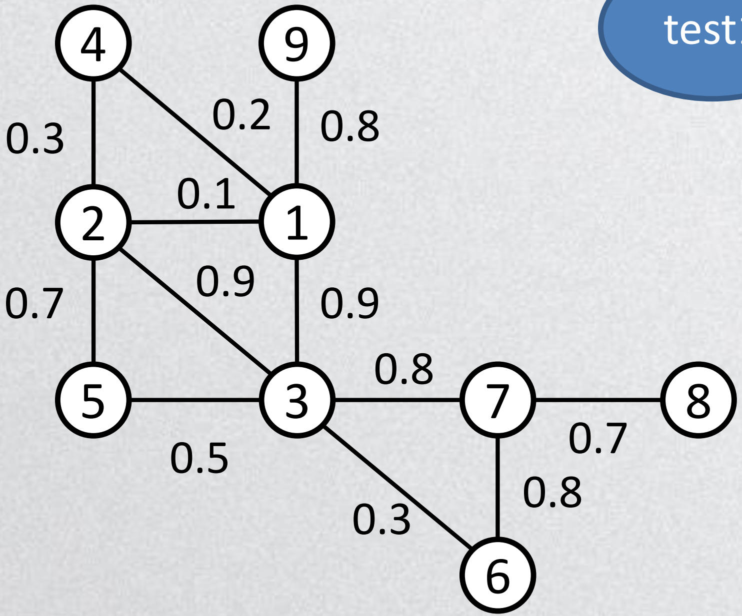
we believe that the difference of those effects between the case $t_v = 0$ and $t_v > 0$ is negligible for small p

思考：如果边上的概率不都等于$p$呢？

# 影响力最大化1：degree discount

运用公式$d_v - 2t_v - (d_v - t_v)t_v p$

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$




test1

d1=1+2.9

d2=1+2.0

**d3=1+3.4**

max

d4=1+0.5

d5=1+1.2

d6=1+1.1

d7=1+2.3

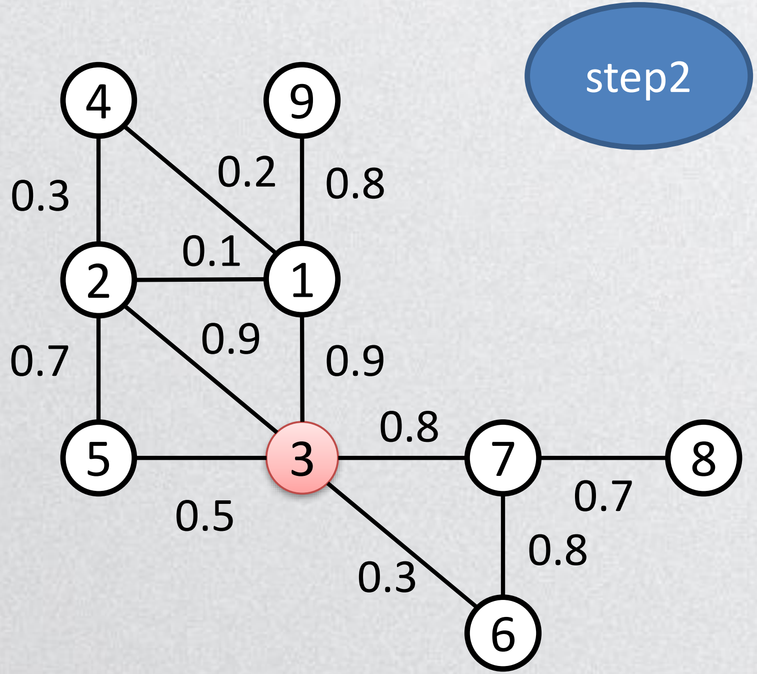
d8=1+0.7

d9=1+0.8

# 影响力最大化1：degree discount

运用公式$d_v - 2t_v - (d_v - t_v)t_v p$

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$




S={3} ,3的邻居有1，2，5，6，7

d1=1+2.9　3.9 – 2*1 – (3.9-1)*1*0.9=-0.71→0

d2=1+2.0　3.0 – 2*1 – (3.0-1)*1*0.9=-0.8 →0

d3=1+3.4

d4=1+0.5

d5=1+1.2　2.2 – 2*1 – (2.2-1)*1*0.5=-0.4 →0

d6=1+1.1　2.1 – 2*1 – (2.1-1)*1*0.3=-0.23 →0

d7=1+2.3　3.3 – 2*1 – (3.3-1)*1*0.8=-0.54 →0

d8=1+0.7

d9=1+0.8

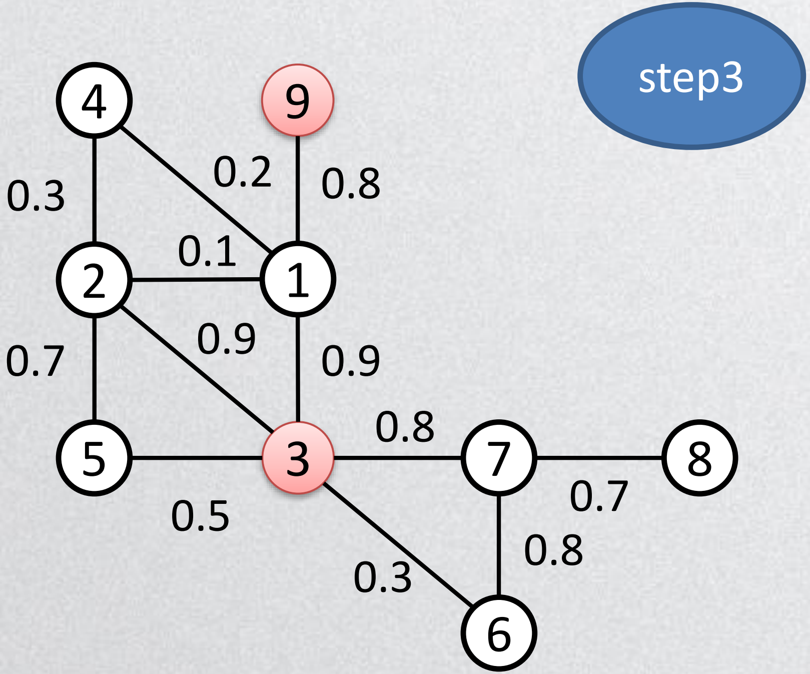
max

出现负值可以归零

# 影响力最大化1：degree discount

运用公式$d_v - 2t_v - (d_v - t_v)t_v p$

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$




S={3，9} ,9的邻居有1

d1=0　0-2*2 – (0-2)*2*0.8 = -0.8 →0

d2=0

d3=1+3.4

d4=1+0.5

d5=0

d6=0

d7=0

d8=1+0.7　max

d9=1+0.8

出现负值可以归零

如果要继续选下一个就是8

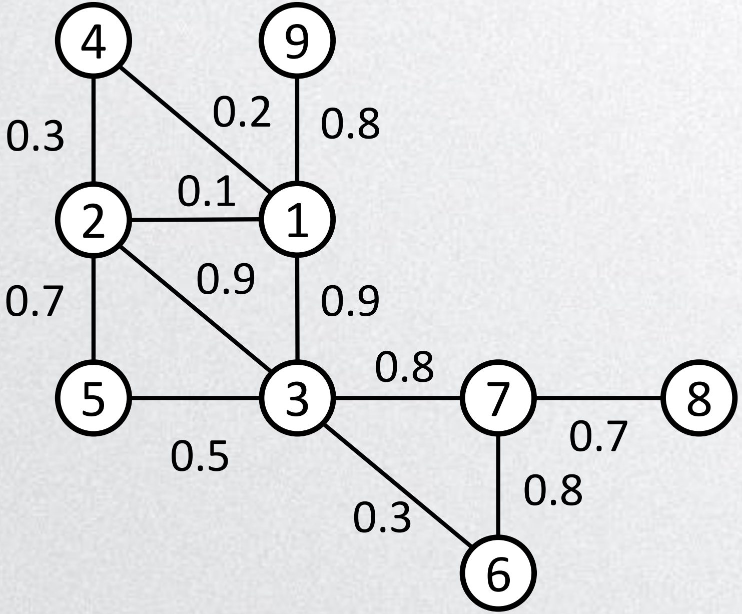
# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 如下示例：基于Degree Discount算法思想，在下图中找出$k=2$时，使影响范围$\sigma(S)$最大的种子集合$S$

- 前文Degree Discount算法
- 针对的是大型网络
- $d_v$比较大
- P比较小

可按照近似算法计算

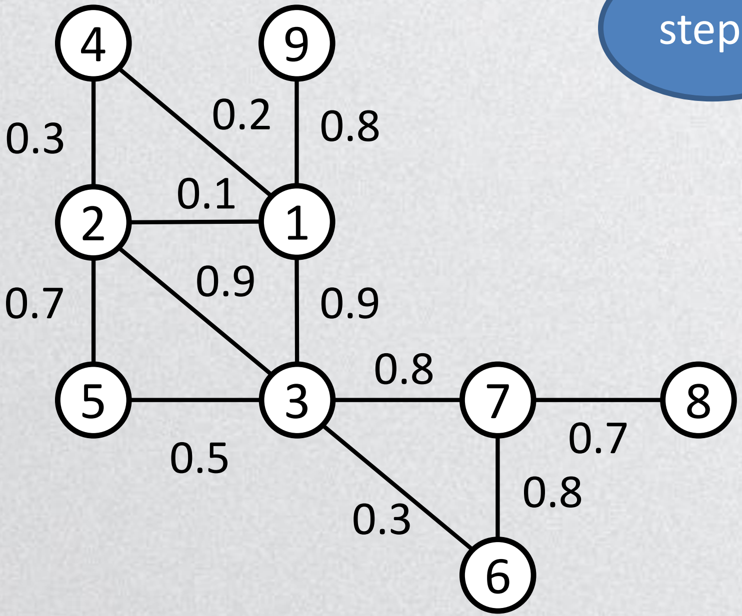



- 本实例
- $d_v$比较小
- P比较大

按照精确算法计算

# 影响力最大化1：degree discount

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$




step1

pp1=1.0, d1=1+2.9, dd1=pp1*d1=3.9

pp2=1.0, d2=1+2.0, dd2=pp2*d2=3.0

**pp3=1.0, d3=1+3.4, dd3=pp3*d3=4.4** max

pp4=1.0, d4=1+0.5, dd4=pp4*d4=1.5

pp5=1.0, d5=1+1.2, dd5=pp5*d5=2.2

pp6=1.0, d6=1+1.1, dd6=pp6*d6=2.1

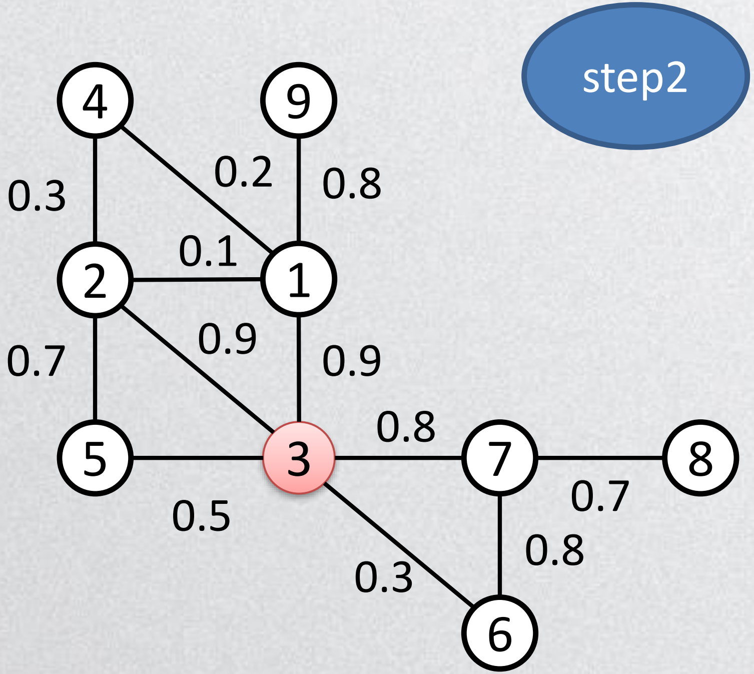
pp7=1.0, d7=1+2.3, dd7=pp7*d7=3.3

pp8=1.0, d8=1+0.7, dd8=pp8*d8=1.7

pp9=1.0, d9=1+0.8, dd9=pp9*d9=1.8

# 影响力最大化1：degree discount

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$




S={3} ,3的邻居有1，2，5，6，7

pp1*=(1-0.9)=0.1，d1=3.9-0.9=3, dd1=0.3

pp2*=(1-0.9), d2=3.0-0.9=2.1, dd2=0.21

pp3=1.0, d3=1+3.4, dd3=pp3*d3=4.4

pp4=1.0, d4=1+0.5, dd4=pp4*d4=1.5

pp5*=(1-0.5), d5=2.2-0.5=1.7, dd5=0.85

pp6*=(1-0.3), d6=2.1-0.3=1.8, dd6=1.26

pp7*=(1-0.8), d7=3.3-0.8=2.5, dd7=0.5

pp8=1.0, d8=1+0.7, dd8=pp8*d8=1.7

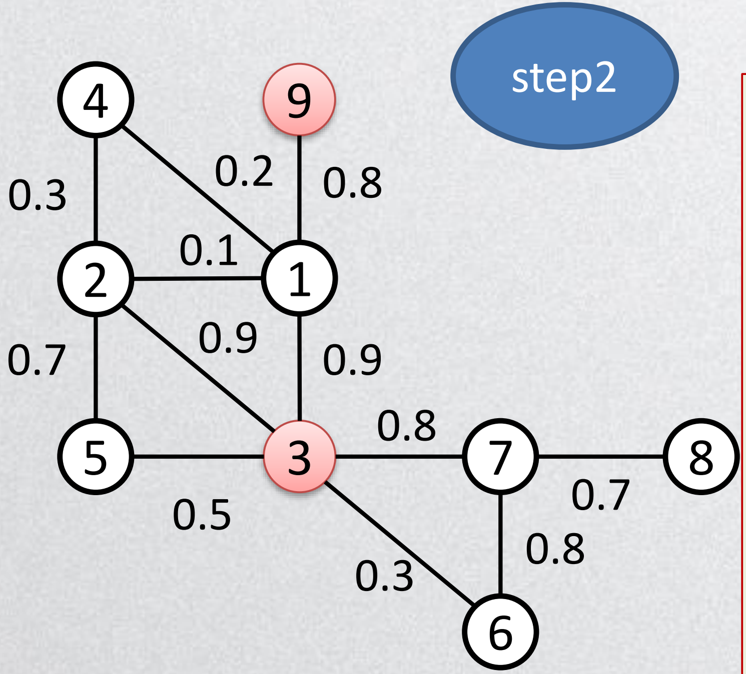
**pp9=1.0, d9=1+0.8, dd9=pp9*d9=1.8**

max

# 影响力最大化1：degree discount

- Degree Discount算法
  - 基于Degree Discount算法思想，在下图中找出$k = 2$时，使影响范围$\sigma(S)$最大的种子集合$S$

S={3，9} ,9的邻居有1




pp1*=0.1*(1-0.8)=0.02，d1=3-0.8=2.2, dd1=0.044

pp2*=(1-0.9), d2=3.0-0.9=2.1, dd2=0.21

pp3=1.0, d3=1+3.4, dd3=pp3*d3=4.4

pp4=1.0, d4=1+0.5, dd4=pp4*d4=1.5

pp5*=(1-0.5), d5=2.2-0.5=1.7, dd5=0.85

pp6*=(1-0.3), d6=2.1-0.3=1.8, dd6=1.26

pp7*=(1-0.8), d7=3.3-0.8=2.5, dd7=0.5

pp8=1.0, d8=1+0.7, dd8=pp8*d8=1.7

pp9=1.0, d9=1+0.8, dd9=pp9*d9=1.8

# 影响力最大化1：degree discount

# 影响力最大化1：degree discount

- 扩展阅读
  - 课上讲的算法仅是最基本的算法
  - 如有兴趣扩展阅读，请参考
  - http://iir.ruc.edu.cn/~fanj/talks/im-tsinghua.pdf