

The Friendship Paradox

Social Networks Analysis and Graph Algorithms

Prof. Carlos Castillo — <https://chato.cl/teach>



**Universitat
Pompeu Fabra**
Barcelona

Contents

- Sampling nodes and edges
- Average degree of friends

Sources

- A. L. Barabási (2016). Network Science – Chapter 04
- F. Menczer, S. Fortunato, C. A. Davis (2020). A First Course in Network Science – Chapter 03
- URLs cited in the footer of specific slides

Sampling a random node

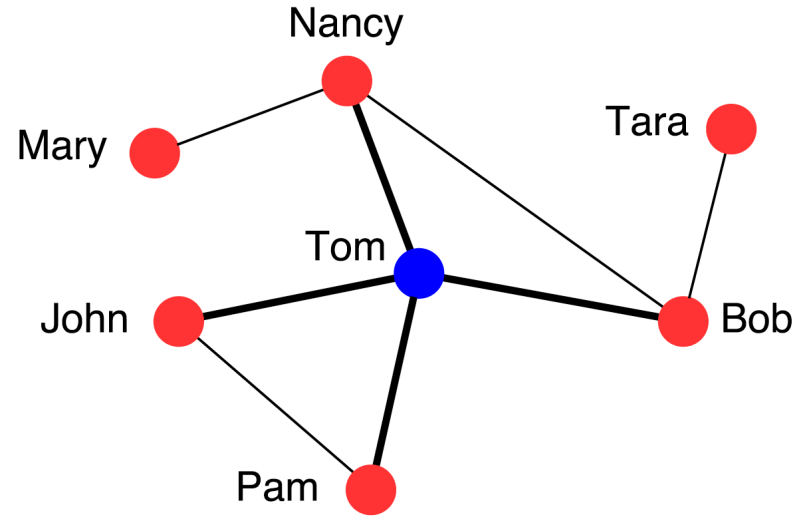
vs

**sampling at random one of the two nodes
attached to a random edge**

Exercise

Numerical calculation of friendship paradox

- What is the probability of selecting Tom **if we select a random node**?
- What is the probability of selecting Tom **if we select a random edge and then randomly one of the two nodes attached to it**?

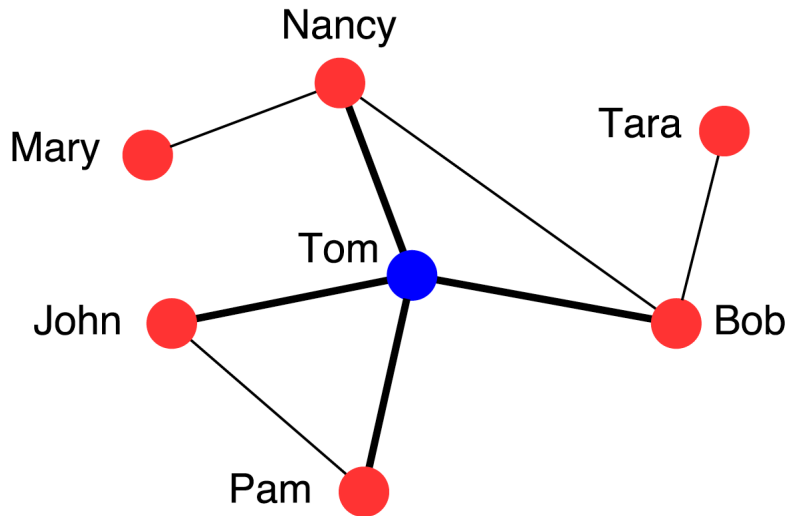


Pin board: <https://upfbarcelona.padlet.org/chato/ocpl5n14i8hrkr4i>



Sampling a random node
vs
sampling a random friend
of a random node

Average degree of friends



- Average degree
$$(1 + 3 + 3 + 1 + 4 + 2 + 2)/7 = 16/7 \approx 2.29$$
- Average degree of friends of:
 - Mary: 3
 - Nancy: $(1+4+3)/3 = 8/3$
 - Tara: 3
 - Bob: $(1+3+4)/3 = 8/3$
 - Tom: $(3+3+2+2)/4 = 10/4$
 - John: $(4+2)/2 = 3$
 - Pam: $(4+2)/2 = 3$
 - Average degree of friends ≈ 2.83 (> 2.29)

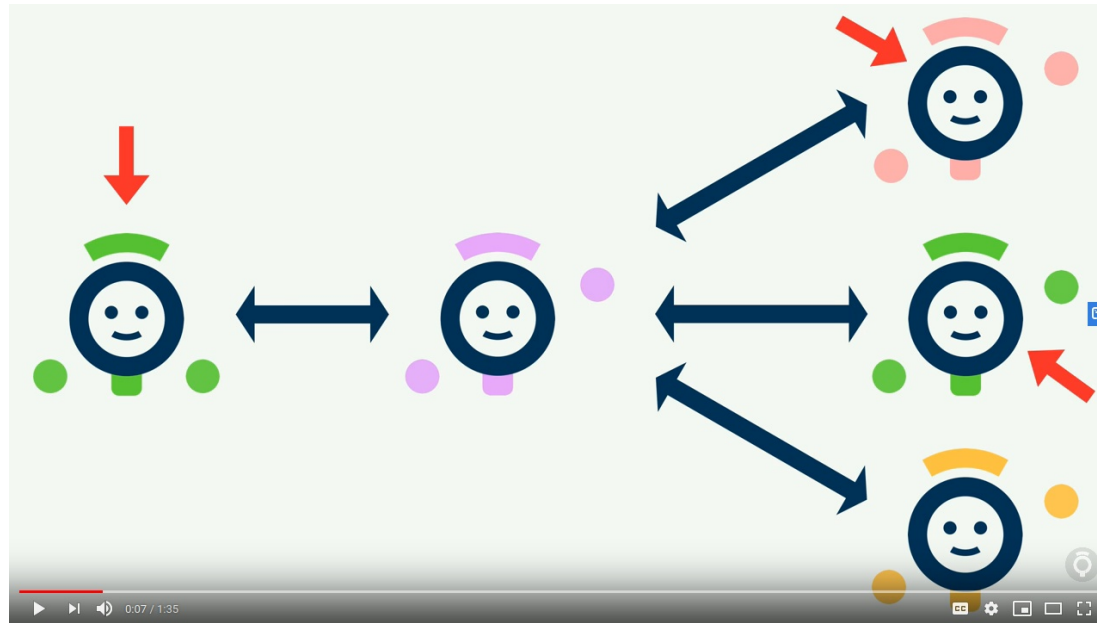
The friendship paradox

- Take a random person x ; what is the expected degree of this person? $\langle k \rangle$
- Take a random person x , now pick one of x 's neighbors, let's say y ; what is the expected degree of y ?

It is not $\langle k \rangle$

- This “paradox” is a useful:
 - As a marketing strategy: if u invites a friend v to buy/use a product, it is likely that v has many friends, and hence it is relevant for marketing that v buys/use the product
 - As a vaccination strategy: instead of offering a vaccine to random people, ask them to name a friend, offer the vaccine to those people, who will have larger degree

Sampling bias and the friendship paradox (1'35'')



<https://www.youtube.com/watch?v=httLvVufAYs>

Imagine you're at a random airport on earth

- Is it more likely to be ...
a large airport or a small airport?
- If you take a random flight out of it ...
will it go to a large airport or a small airport?

An example of friendship paradox

- Pick a random airport on Earth
 - Most likely it will be a small airport
- However, no matter how small it is, it **will** have flights to big airports
- On average those airports will have much larger degree



Time	Flight	Airline	Destination	Gate	Exp.	Remarks
11:00	KA 376	DRAGONAIR	Hong Kong	4		Chk-in closed
12:25	DG 7792	tigerair	Singapore	1		On Time
12:25	QR 931	QATAR	Doha, Qatar	5		On Time
17:40	EK 339	Emirates	Dubai	5		On Time
00:50	OZ 708	ASIANA AIRLINES	Seoul Incheon	5		On Time
07:05	5J 150	JAL	Hong Kong	1		On Time
07:20	DG 7924	tigerair	Hong Kong	1		On Time
08:00	DG 7792	tigerair	Singapore	1		On Time
12:10	5J 537	JAL	Singapore	1		On Time
12:25	QR 931	QATAR	Doha, Qatar	5		On Time

Exercise [B. 2016, Ex. 4.10.2]: "Friendship Paradox"

- If random variable K represents the degree of a randomly chosen node, we denote as p_k the probability that a randomly chosen node has degree k
 - $p_k = \Pr(K=k)$ Note that for simplicity we always denote by $\langle k \rangle$ what we should have named $\langle K \rangle$
- Random variable K_F will represent the degree of a randomly chosen neighbor ("friend") of a randomly chosen node; we will denote by q_k the probability that a randomly chosen neighbor of a randomly chosen node has degree k
$$q_k = \Pr(K_F=k)$$
- The formula is: $q_k = C k p_k$ where C is a normalization factor
 - (a) Find C (hint: sum of q_k must be 1)

Exercise [B. 2016, Ex. 4.10.2]: "Friendship Paradox"

Random variable K_F is the degree of a randomly chosen neighbor of a randomly chosen node; we denote by q_k the probability that a randomly chosen neighbor of a randomly chosen node has degree k

$$q_k = \Pr(K_F=k) = C k p_k$$

(b) Find the expectation $\langle K_F \rangle$

Hints: $E[X] = \sum_x x \cdot P(X = x)$ $E[X^2] = \sum_x x^2 \cdot P(X = x)$

Exercise [B. 2016, Ex. 4.10.2]: "Friendship Paradox"

For the scale-free network described below:

(c) Compute $\langle K_F \rangle$: the expected number of friends of a randomly chosen neighbor of a randomly chosen node

(d) Compare with $\langle k \rangle$: the expected number of friends of a randomly chosen node

$$N = 10000$$

$$\gamma = 2.3$$

$$k_{\min} = 1$$

$$k_{\max} = 1000$$

You can use this formula for the **moments** ($\langle k \rangle$, $\langle k^2 \rangle$, $\langle k^3 \rangle$, ...) of the degree distribution in a scale-free network:

$$\langle k^n \rangle = (\gamma - 1) k_{\min}^{\gamma-1} \frac{\left(k_{\max}^{n-\gamma+1} - k_{\min}^{n-\gamma+1} \right)}{n - \gamma + 1}$$

Code

```
def degree_moment(kmin, kmax, moment, gamma):  
    C = (gamma-1.0)*(kmin**(gamma-1.0))  
    numerator = (kmax**(moment-gamma+1.0) - kmin**(moment-gamma+1.0))  
    denominator = (moment-gamma+1.0)  
    return C * numerator / denominator
```

```
kavg = degree_moment(kmin=1, kmax=1000, moment=1, gamma=2.3)  
print(kavg)
```

3.787798988222529

```
ksqavg = degree_moment(kmin=1, kmax=1000, moment=2, gamma=2.3)  
print(ksqavg)
```

231.94329076177414

```
print(ksqavg / kavg)
```

61.23431879119234

Summary

Summary

- Your friends have more friends than you

$$\langle K_F \rangle > \langle k \rangle$$

- This can be quite strong in scale-free networks

Practice on your own

- Draw a small graph, and sample from that graph until you're convinced $\langle K_F \rangle > \langle k \rangle$