On the friendship paradox

Elizaveta Rebrova

ORF387, February 2024

The friendship paradox informally says that

> "Most people have fewer friends than their friends have, on average.".

To explain it, we need to make precise several things: How is this "average" computed? How should we distinguish between "people on average" and their friends?

The general idea is that we can define how we sample (take) a random member of a network in two different ways: (1) as "a random person" and (2) as "a random friend" of someone else. Then, we can compare the expected number of friends of a randomly selected vertex in these two cases.

A random person: It is intuitive to define a random person in a network $G = (V, E)$ as a vertex of a network chosen uniformly at random from the set V. So, $v \in V$ is chosen as a random person with probability $1/|V|$, where $|V|$ denotes the number of elements in the set V. Then, the expected number of friends of a random person is

$$\begin{aligned}&\mathbb{E}(\text{number of friends of a random person})\\ &\quad = \sum_{v \in V} \mathbb{P}(\text{choose } v \text{ as a random person}) \cdot [\text{number of friends of } v]\\ &\quad = \sum_{v \in V} \frac{1}{|V|} \cdot \deg(v) = \frac{1}{|V|} \sum_{v \in V} \deg(v) = \frac{2|E|}{|V|},\end{aligned}$$

the last step holds as summing all vertex degrees means counting each edge in the network twice (from both sides).

A random friend: Now, what is the probability that a particular vertex $v \in V$ is named to be "a random friend"? We modeled this in class in the following wa:

1. Pick a random friendship (edge of the graph)
2. Pick a random member of this friendsdhip (one of the two with probability 1/2).

A vertex chosen on step 2 is the selected "random friend". Under this model, the probability that $v \in V$ is chosen as a random friend is

$$\mathbb{P}(\text{pick a friendship that involves } v) \cdot \mathbb{P}(\text{pick } v \text{ but not the other friend}) = \frac{deg(v)}{|E|} \cdot \frac{1}{2}.$$

Several observation regarding this probability distribution:

- It is a well-defined probability distribution: check that all the individual probabilities sum up to 1!
- It makes sense: the larger degree of a vertex (=number of friends of a person) the higher is the probability to sample this person as someone's friend.
- Another way to explain these probabilities: When is v "a random friend"? It is when someone else called them "their friend". Ok, let's ask everyone in the network who their friends are, make a list of mentions and define the probability that "v is a random friend" to be a fraction of mentions of v in this total list of friends. Note that v will be mentioned by every friend of v (once), so there will be $\deg(v)$ mentions in the list. How many records are there in the list? Everyone mentioned all their friends, so, every friendship was called twice, and, in total, there should be $2|E|$ records.

Then, the expected number of friends of a random friend is

$$\begin{aligned}
&\mathbb{E}(\text{number of friends of a random friend}) \\
&\quad = \sum_{v \in V} \mathbb{P}(\text{choose } v \text{ as a random friend}) \cdot [\text{number of friends of } v] \\
&\quad = \sum_{v \in V} \frac{\deg(v)}{2|E|} \cdot \deg(v) = \frac{1}{2|E|} \sum_{v \in V} \deg^2(v).
\end{aligned}$$

Theorem 1. *We claim that*

$$\mathbb{E}(\textit{number of friends of a random person}) \leq \mathbb{E}(\textit{number of friends of a random friend})$$

and the equality is achieved if an only if the graph is regular (all vertices have the same number of neighbors), which is super rare for a social graph.

Proof. From the computations above, for $\mu := 2|E|/|V|$, we need to show that

$$\mu \leq \frac{1}{2|E|} \sum_{v \in V} \deg^2(v) = \frac{1}{\mu} \cdot \frac{1}{|V|} \sum_{v \in V} \deg^2(v).$$

Note that the last averaged sum is the second moment of the number of friends of the random person, and μ is its expectation. So,

$$\frac{1}{|V|} \sum_{v \in V} \deg^2(v) = \mu^2 + \sigma^2, \text{ where } \sigma^2 := \frac{1}{|V|} \sum_{v \in V} (\mu - \deg(v))^2 \geq 0.$$

Exercise: check that the last statement holds by definition of μ without referring to the second moment-expectation-variance relations. In conclusion,

$$\frac{1}{\mu} \cdot \frac{1}{|V|} \sum_{v \in V} \deg^2(v) = \frac{1}{\mu}(\mu^2 + \sigma^2) = \mu + \mu/\sigma^2 \geq \mu$$

and the equality is achieved if and only if $\sigma^2 = 0$, so, every term in the definition of σ^2 is equal to zero, so, if the degrees of all the vertices are equal to μ and equal to each other. $\square$

Why is the friendship paradox interesting? Here are several reasons:

- It shows us how subtle is the notion of picking something "at random" from the environment, it is not always just about picking a random member from the set of all members (since interactions matter!)
- The "random friend sampling" idea is an easy way to *sample with preference to the vertices with higher degree centrality*, so, it is useful beyond explaining the friendship paradox. For example, some studies used it to forecast and slow the course of epidemics: the higher centrality vertices are individuals to immunize or monitor for infection, and the random selection of several "random friends" allows to identify such vertices without complex computation of the centrality of all nodes in the network.
- There are also "generalized friendship paradoxes" that give similar statements about other characteristics of the average friends, such as, popularity, productivity, and even happiness. This touches on generally important idea that the things from one's point of view don't have to ideally reflect reality, and it is not right (and potentially feeds the impostor syndrome) to base your assumptions of popularity on the popularity of your friends. See also HW2 problem 3.

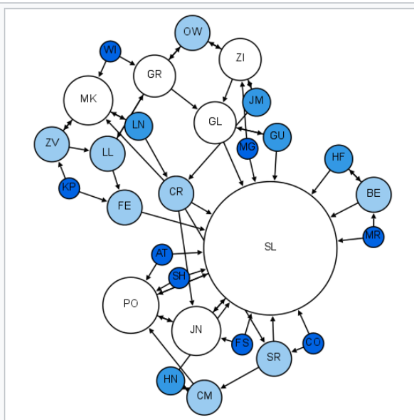

Diagram of a social network of 7-8 year old children, mapped by asking each child to indicate two others they would like to sit next to in class. The majority of children have fewer connections than the average of those they are connected to.

The picture is taken from the wikipedia article `https://en.wikipedia.org/wiki/Friendship_paradox`, this article is also a good source for many related references.