Small worlds and degrees of separation

Seminars in Social Networks and Markets

The first notion of “small world”

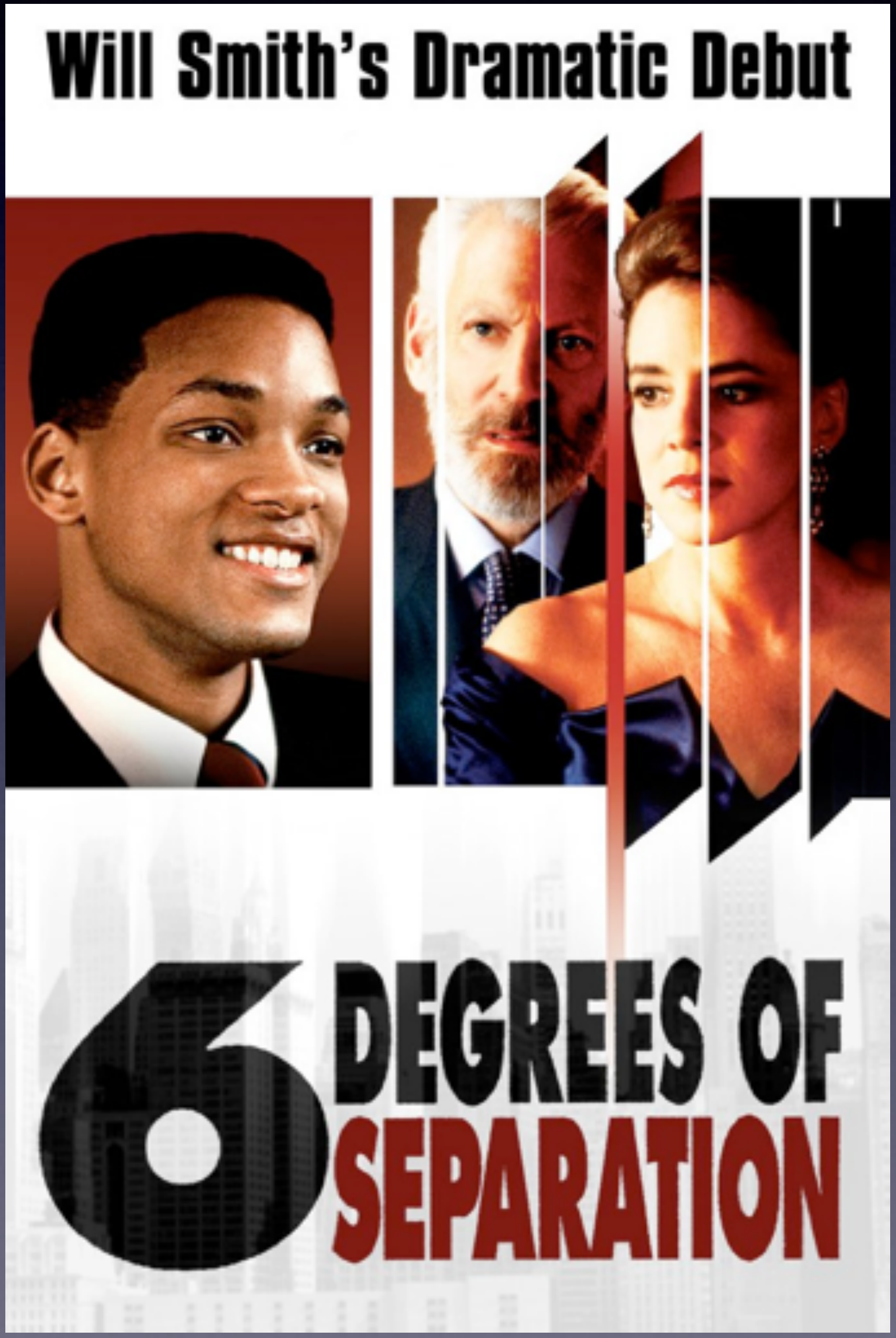
- *“The population of the Earth is closer together now than they have ever been before”*

- At most 6 links to get from you to anyone else on the planet?

- *Chains (Láncszemek)*, 1929
Frigyes Karinthy (1887–1938): Hungarian author

In popular culture

- Inspired "*Six Degrees of Separation*"
  - play by J. Guare, 1990
  - movie by F. Schepisi, 1993
  - TV series, 2006



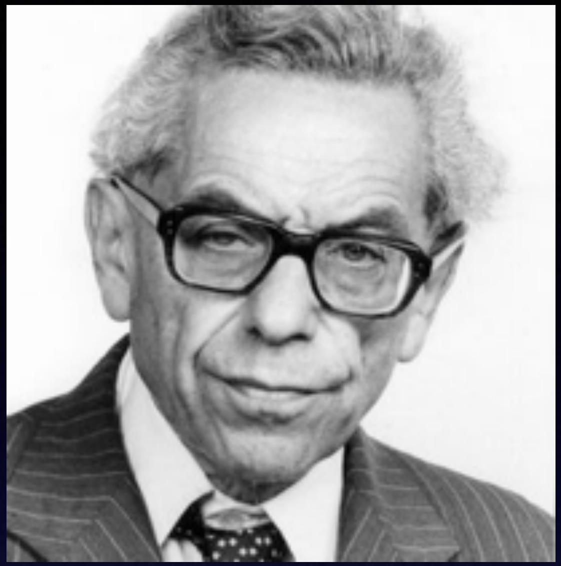
Kevin Bacon game

- www.oracleofbacon.org
- Actor collaboration graph
- How many steps to reach Kevin Bacon?
- Example:

Will Smith has a **Bacon number** of 2:

Will Smith
↓
Enemy of the State (1998)
↓
Frank Medrano
↓
Sleepers (1996)
↓
Kevin Bacon

Erdős numbers



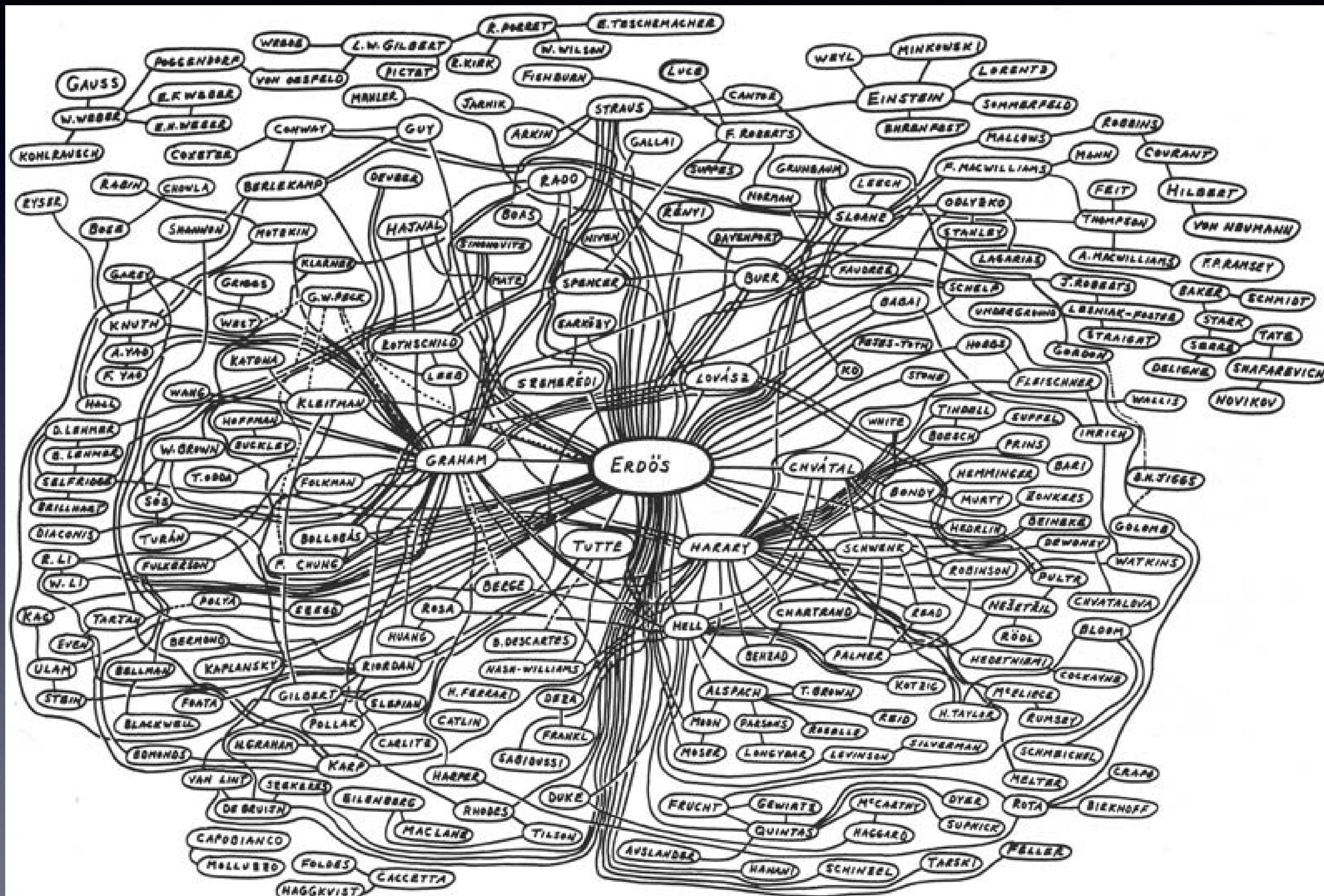
- Paul Erdős (1913–1966), Hungarian mathematician
- Mathematics collaboration graph
- How many steps to reach Paul Erdős?
- Example:

My **Erdős number** is 3:

Vincenzo Bonifaci
↓
Kurt Mehlhorn
↓
Dieter Kratsch
↓
Paul Erdős

- www.ams.org/mathscinet/collaborationDistance.html

Giant component



The small-world experiment

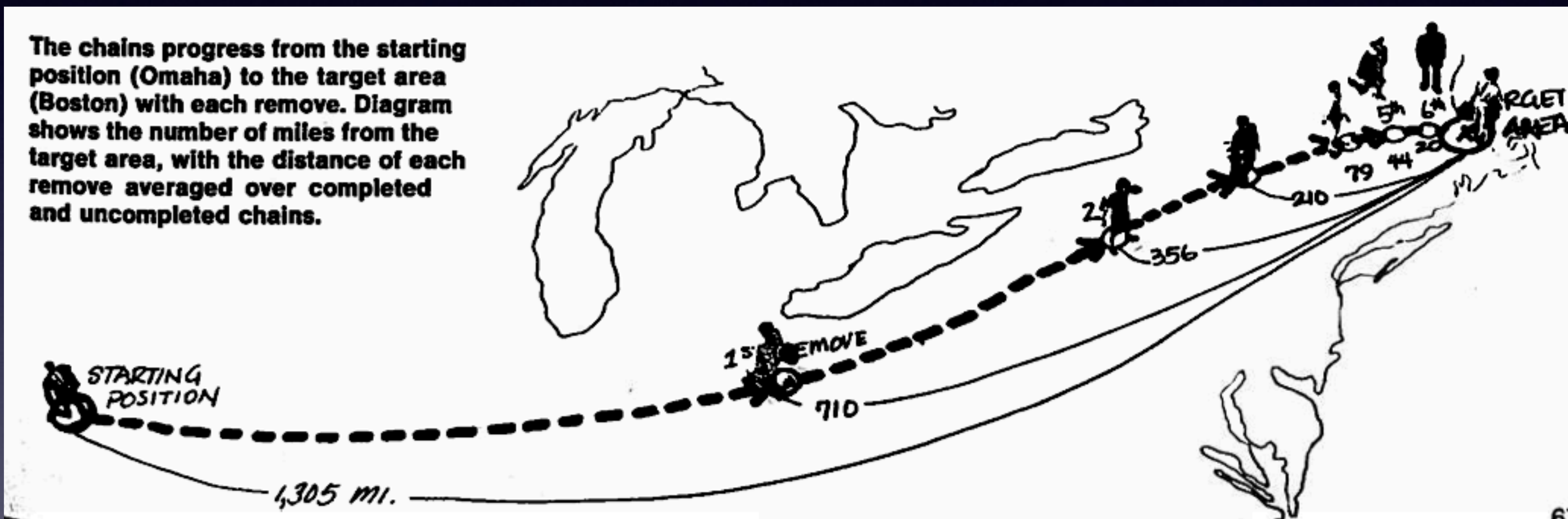
- Study by sociologist Stanley Milgram in 1969 (with J.Travers)

- Milgram's question*:*
  "Given two individuals selected randomly from the population, what is the probability the minimum number of intermediaries required to link them is 0, 1, 2, …, k?"

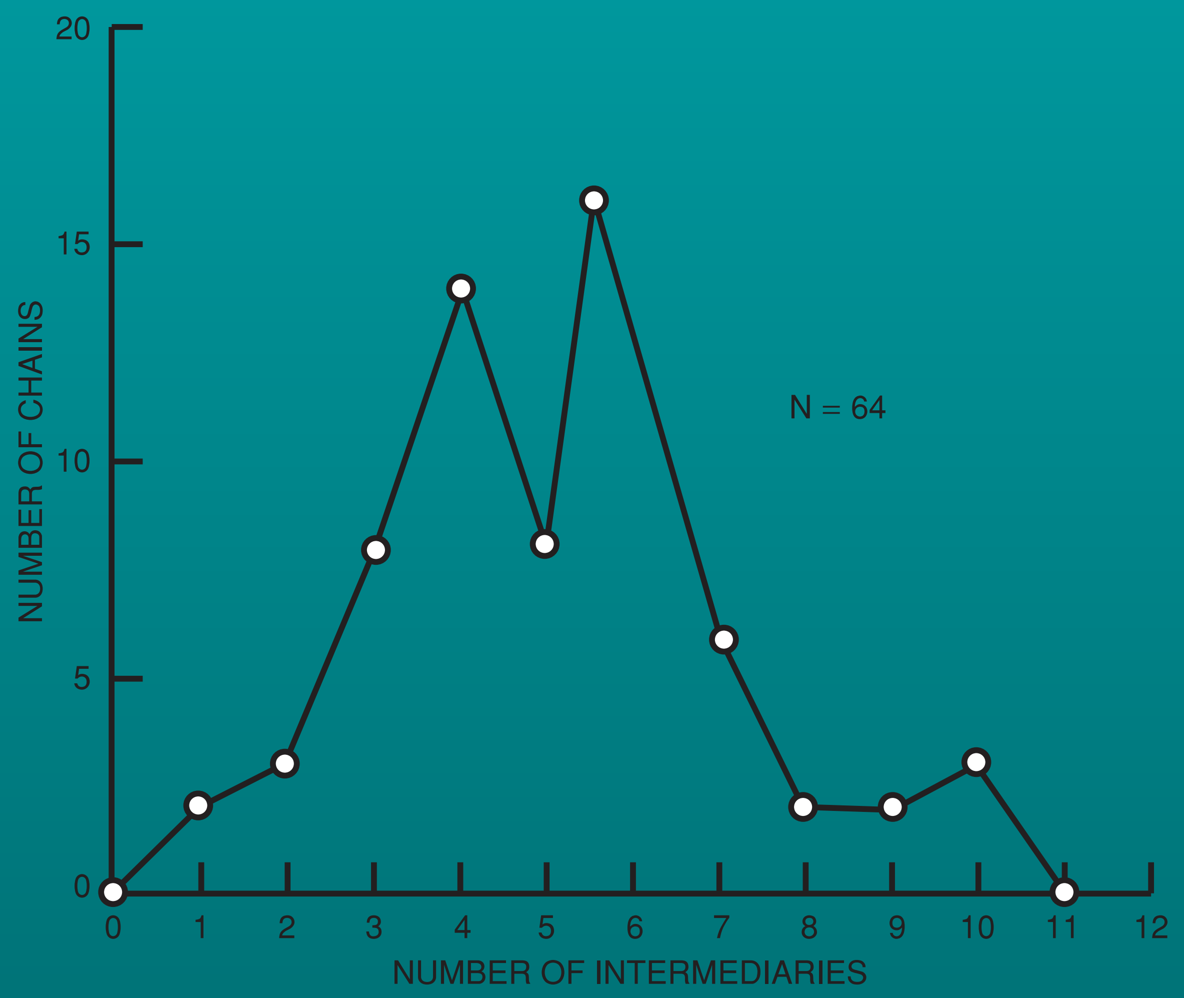
The small-world experiment

- 296 people asked to dispatch a parcel to a single individual (target)
- Target: a Boston stockholder
- Starting population:
  - 100 random Boston inhabitants
  - 100 random Nebraska stockholders
  - 96 random Nebraska inhabitants
- **Rule of the game:** the parcel can only be sent directly to a personal acquaintance ("first-name acquaintance")

The small-world experiment



The small-world experiment



- Average distance was 6.2 (5.4, 6.4, 6.7 for group 1, 2, and 3 resp.)
- Distance 6.7 = 5.7 degrees of separation (origin of the "six degrees" notion)
- 29% of the parcels reached the target

Still relevant?

- What about today's online social networks?
  - Facebook:
    - 4.74 average distance in 2011
    - 3.57 reported in 2016
  - Twitter: 4.67 average distance in 2010

How do people *find* the next hop?

- The experiment demonstrates two facts:

    1. short paths abound in social networks

    2. that a *decentralized algorithm* exists for effectively finding them

- Problem revisited by Jon Kleinberg in 2000

The graph diameter problem

- The **diameter** of a graph can be computed in time O(n(m+n)) by computing **all** pairwise distances
- It can also be **approximated** in time O(m+n) within a factor 2
- Related (more robust) notion:
  Effective diameter = min # of hops within which 90% of the nodes can reach each other

Clustering coefficient

Measuring triadic closure

- **Clustering coefficient of a node** *u* is the probability that 2 random friends of *u* are friends
- **Global clustering coefficient** is the fraction of length-2 paths that can be extended to a triangle:

$$C(G) = \frac{\#\text{ of closed triplets}}{\#\text{ of connected triplets}}$$

Alternative formulation

$$C(G) = \frac{\#\text{ of closed triplets}}{\#\text{ of connected triplets}}$$

$$C(G) = \frac{3 \cdot \#\text{ of triangles}}{\#\text{ of connected triplets}}$$

since each triangle ABC has 3 triplets:
ABC, BCA, CAB

Social networks have high *C*

Network	C(G)
IMDb actor collaboration network	0.20
Biologists collaboration network	0.09
Email contacts network of a large university	0.16
Internet autonomous system network	0.01

Computing the clustering coefficient

$$C(G) = \frac{\#\text{ of closed triplets}}{\#\text{ of connected triplets}}$$

- Both numerator and denominator can be computed in time $O(n^3)$
- Slow for large networks!
    - Resort to approximation, see Schank & Wagner (2005)