NATIONAL RESEARCH
UNIVERSITY

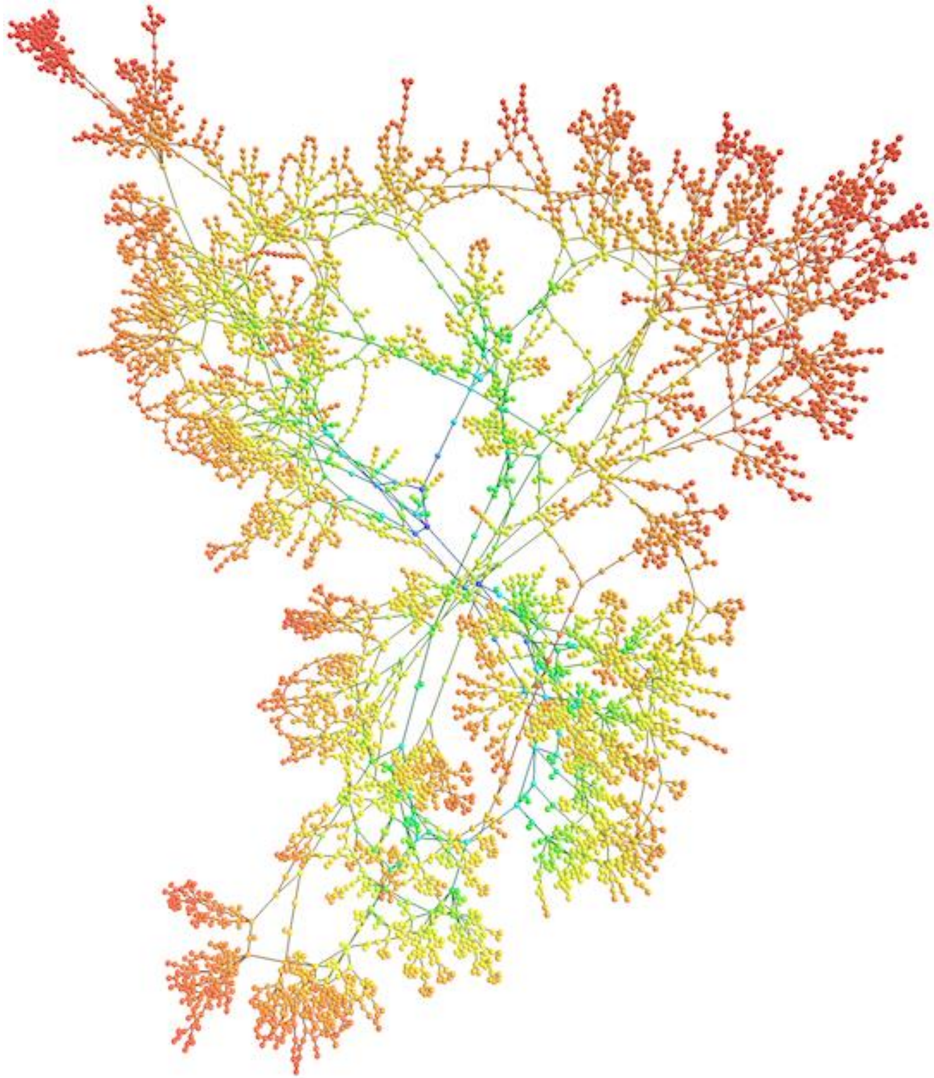
SOCIAL NETWORK DATA ANALYSIS

Ilia Karpov *(*karpovilia@gmail.com*)*

Moscow, 2023

COMPLEX NETWORKS

- not regular, but not random
- non-trivial topology
- universal properties
- everywhere
- complex systems



COMPLEX NETWORKS

- Power law node degree distribution: ”scale-free” networks
- Small diameter and average path length: ”small world” networks
- Hight clustering coefficient: transitivity

POWER LAW

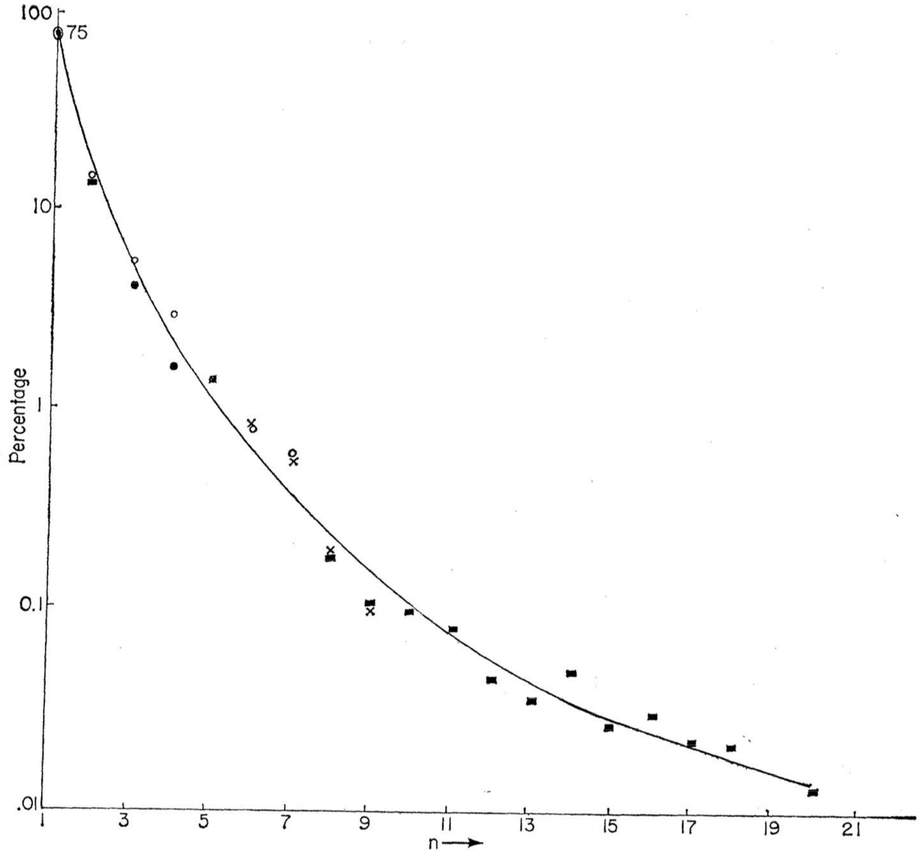
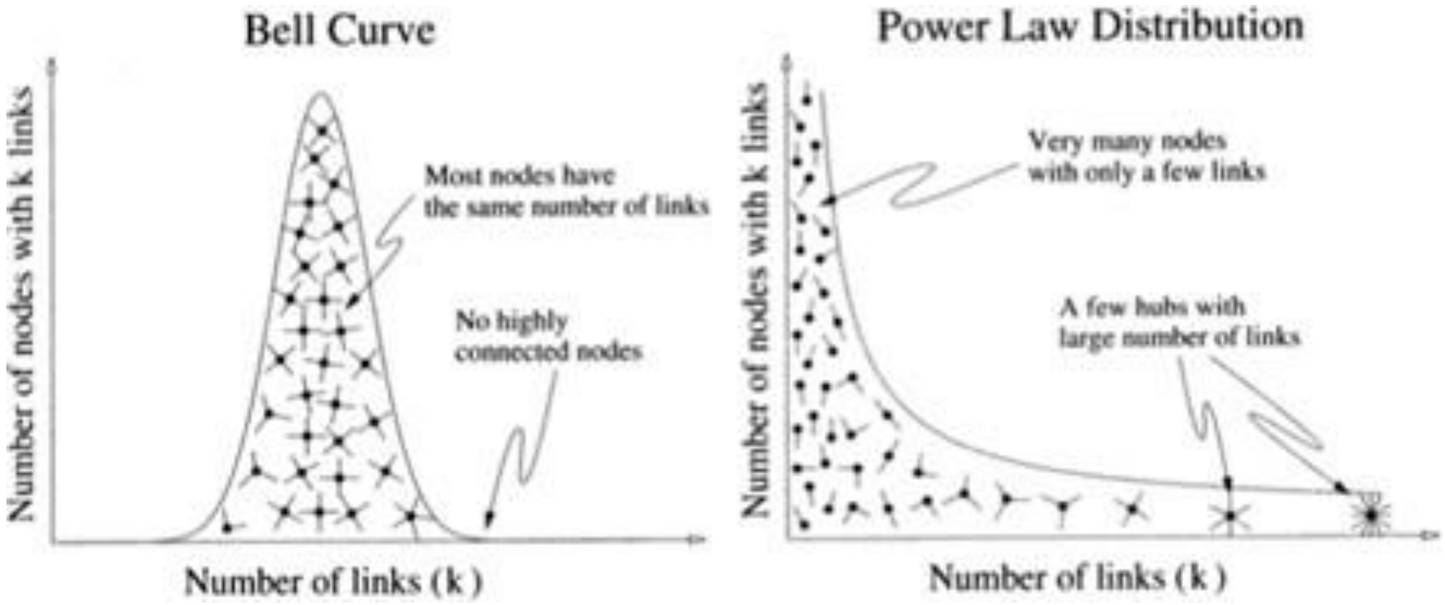


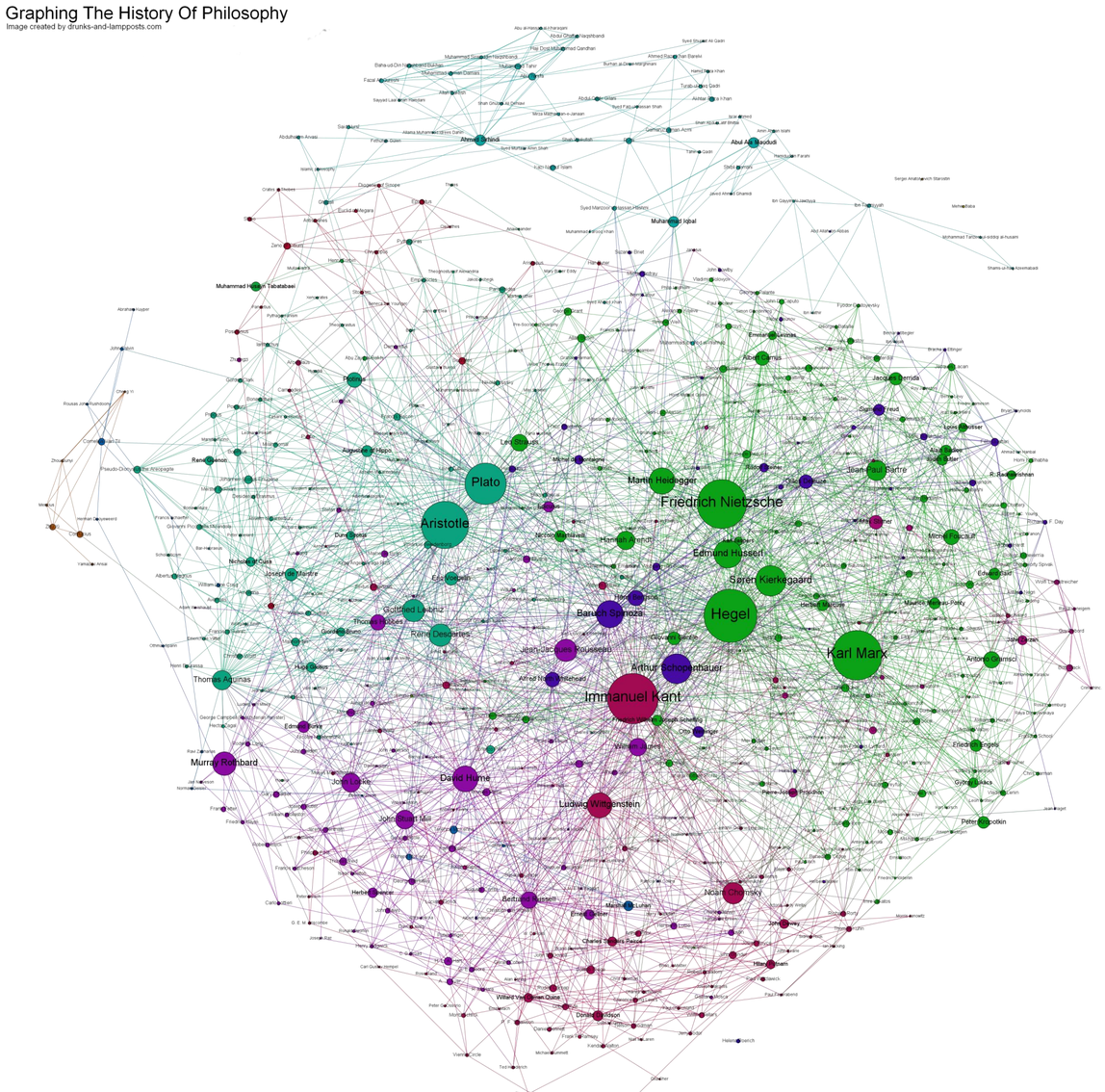
Fig. 2. Percentages (relative to total number of cited papers) of papers cited various numbers (n) of times, for a single year (1961). The data are from Garfield's 1961

POWER LAW

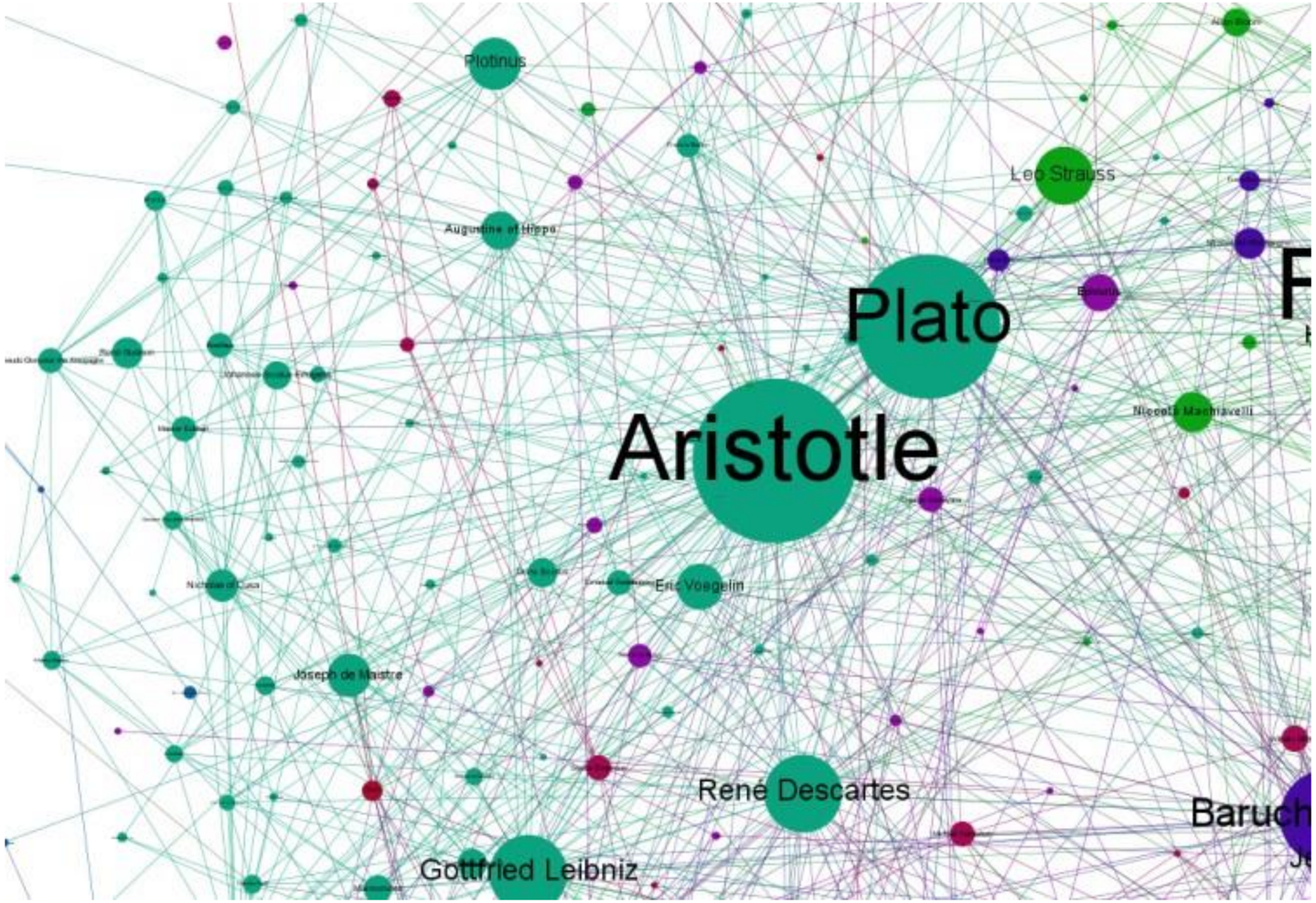

Bell Curve
Number of nodes with k links
Most nodes have the same number of links
No highly connected nodes
Number of links (k)
Power Law Distribution
Number of nodes with k links
Very many nodes with only a few links
A few hubs with large number of links
Number of links (k)

POWER LAW

Lecture 1



POWER LAW


Plotinus
Leo Strauss
Augustine of Hippo
Plato
Aristotle
Niccolò Machiavelli
Eric Voegelin
Joseph de Maistre
René Descartes
Gottfried Leibniz

TRIADS

The Strength of Weak Ties[1]

Mark S. Granovetter
Johns Hopkins University

Analysis of social networks is suggested as a tool for linking micro and macro levels of sociological theory. The procedure is illustrated by elaboration of the macro implications of one aspect of small-scale interaction: the strength of dyadic ties. It is argued that the degree of overlap of two individuals' friendship networks varies directly with the strength of their tie to one another. The impact of this principle on diffusion of influence and information, mobility opportunity, and community organization is explored. Stress is laid on the cohesive power of weak ties. Most network models deal, implicitly, with strong ties, thus confining their applicability to small, well-defined groups. Emphasis on weak ties lends itself to discussion of relations *between* groups and to analysis of segments of social structure not easily defined in terms of primary groups.

- ”The Strength of Weak Ties”, Mark Grannoveter, 1973
- ”Spread of Information through a Population with Socio-Structural Bias. Assumption of Transitivity”, Anatol Rapoport, 1953

TRIADS

- strength of a tie
- high transitivity
- high clustering coefficient

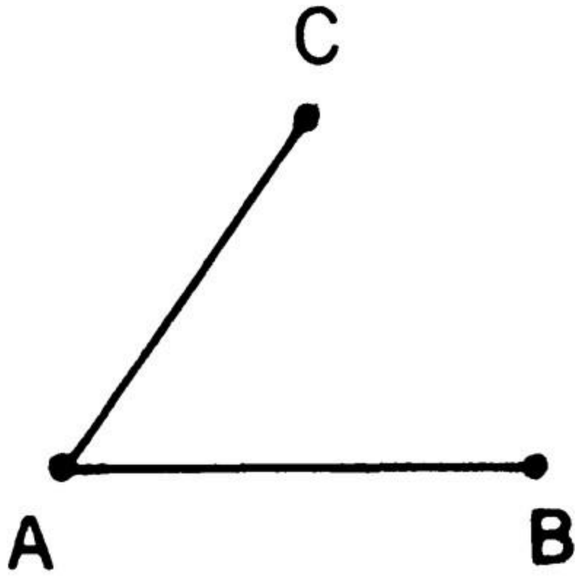


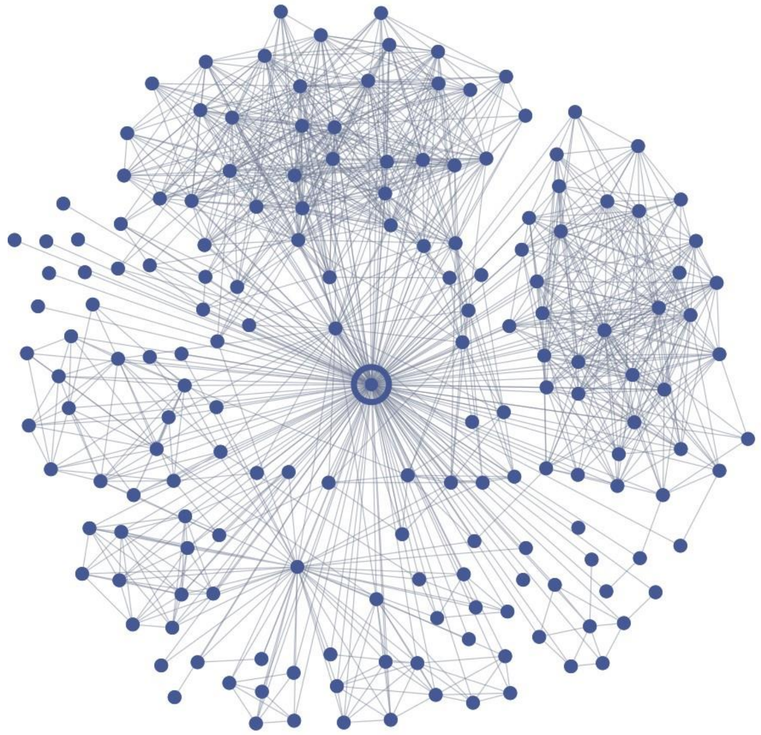
FIG. 1.—Forbidden triad

- If A and B and B and C are strongly linked, the the tie between B and C is always present

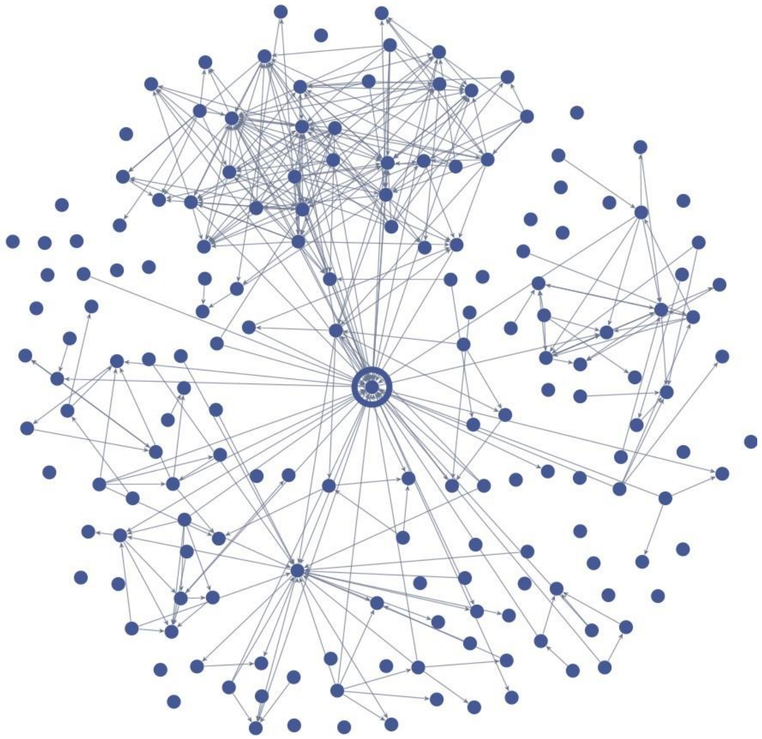
Grannoveter, 1973

HIGH CLUSTERING

All Friends



Maintained Relationships



HIGH CLUSTERING


© Al Satterwhite

An Experimental Study of the Small World Problem*

JEFFREY TRAVERS

Harvard University

AND

STANLEY MILGRAM

The City University of New York

Arbitrarily selected individuals (N=296) in Nebraska and Boston are asked to generate acquaintance chains to a target person in Massachusetts, employing "the small world method" (Milgram, 1967). Sixty-four chains reach the target person. Within this group the mean number of intermediaries between starters and targets is 5.2. Boston starting chains reach the target person with fewer intermediaries than those starting in Nebraska; subpopulations in the Nebraska group do not differ among themselves. The funneling of chains through sociometric "stars" is noted, with 48 per cent of the chains passing through three persons before reaching the target. Applications of the method to studies of large scale social structure are discussed.

- ”The small-world problem”. Stanley Milgram, 1967
- ”An experimental study of the small world problem”, Jeffrey Travers, Stanley Milgram, 1969

STANLEY MILGRAM'S 1967 EXPERIMENT

HOW TO TAKE PART IN THIS STUDY

1. ADD YOUR NAME TO THE ROSTER AT THE BOTTOM OF THIS SHEET, so that the next person who receives this letter will know who it came from.

2. DETACH ONE POSTCARD. FILL IT OUT AND RETURN IT TO HARVARD UNIVERSITY. No stamp is needed. The postcard is very important. It allows us to keep track of the progress of the folder as it moves toward the target person.

3. IF YOU KNOW THE TARGET PERSON ON A PERSONAL BASIS, MAIL THIS FOLDER DIRECTLY TO HIM (HER). Do this only if you have previously met the target person and know each other on a first name basis.

4. IF YOU DO NOT KNOW THE TARGET PERSON ON A PERSONAL BASIS, DO NOT TRY TO CONTACT HIM DIRECTLY. INSTEAD, MAIL THIS FOLDER (POSTCARDS AND ALL) TO A PERSONAL ACQUAINTANCE WHO IS MORE LIKELY THAN YOU TO KNOW THE TARGET PERSON. You may send the folder

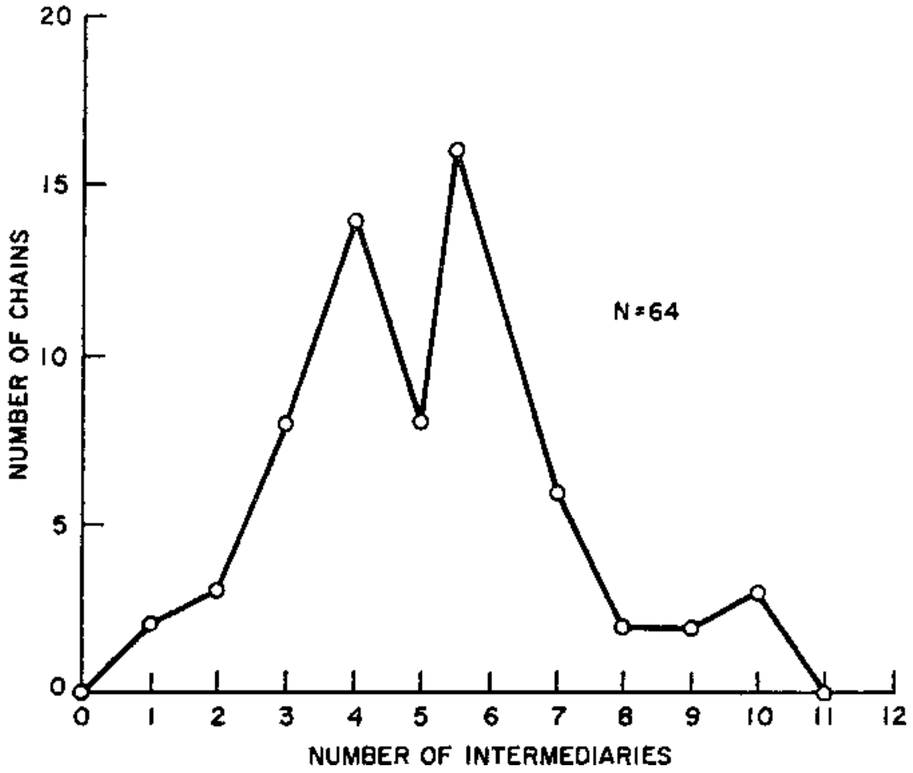
STANLEY MILGRAM'S 1967 EXPERIMENT

- Starting persons:
  - 296 volunteers, 217 sent
  - 196 in Nebraska
  - 100 in Boston
- Target person - Boston stockbroker
- Information given: target name, address, occupation, place of employment, college, hometown

J. Travers, S. Milgram, 1969

STANLEY MILGRAM'S 1967 EXPERIMENT

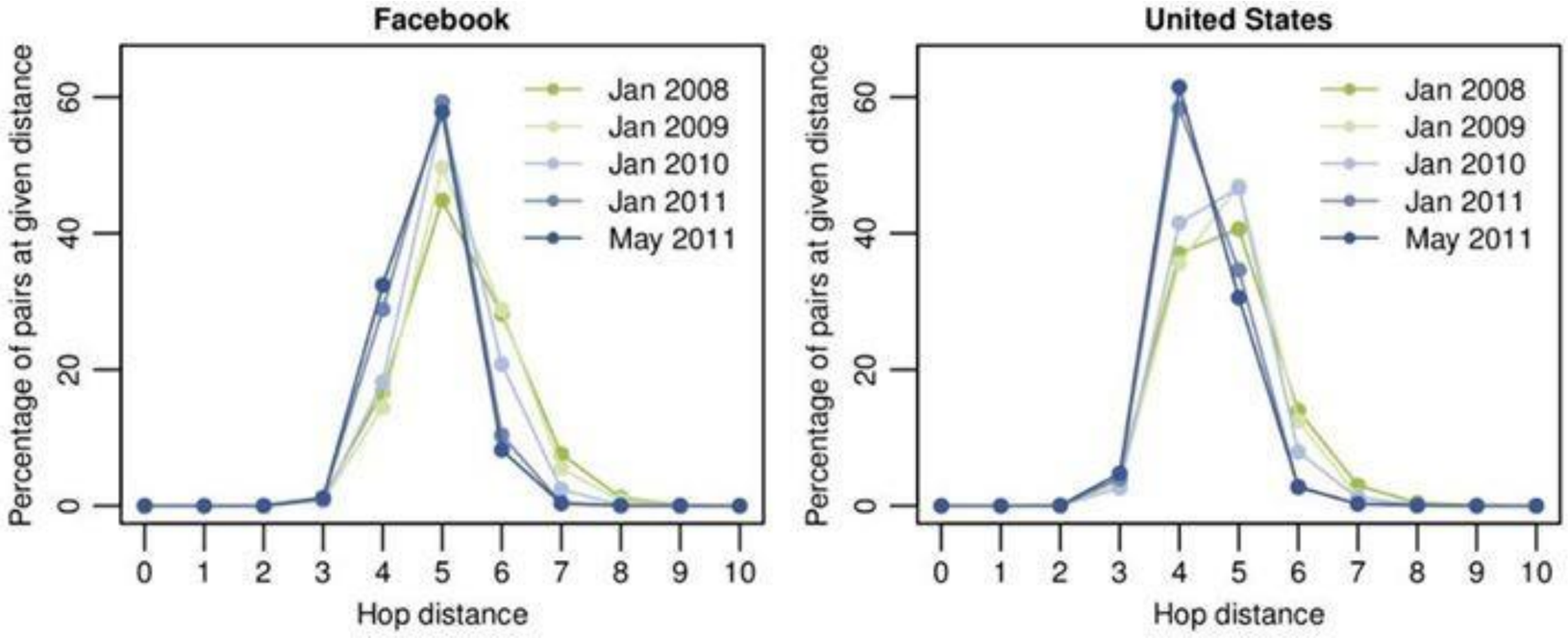
- Reached the target N = 64(29%)
- Average chain length (L) = 5.2
- Channels:
  - hometown (L) = 6.1
  - business contacts (L) = 4.6
  - from Boston (L) = 4.4
  - from Nebraska (L) = 5.7



J. Travers, S. Milgram, 1969

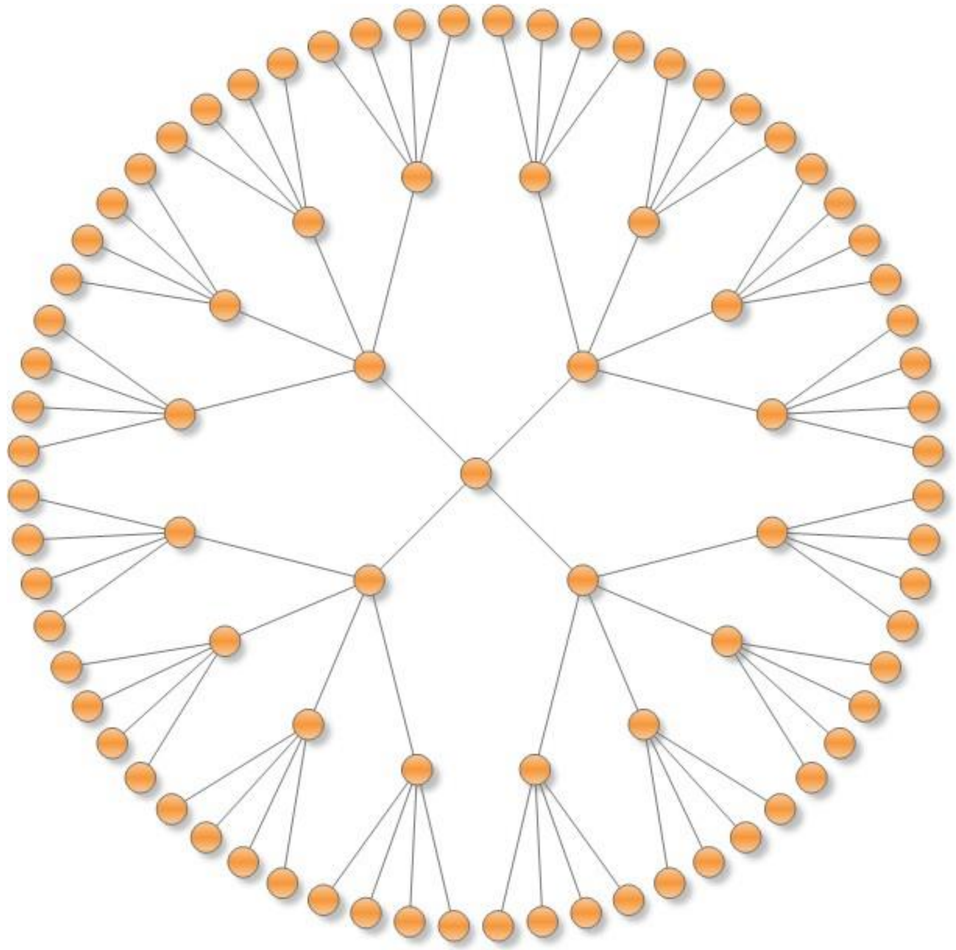
STANLEY MILGRAM'S 1967 EXPERIMENT

- Email graph:
  D. Watts (2001), 48,000 senders, $\langle L \rangle \approx 6$
- MSN Messenger graph:
  J. Lescovec et al (2007), 240mln users, $\langle L \rangle \approx 6.6$
- Facebook graph:
  L. Backstrom et al (2012), 721 mln users, $\langle L \rangle \approx 4.74$



L.Backstrom, 2012

SIMPLE MODEL



An estimate: $z^d = N$, $d = \log N / \log z$

$N \approx 6.7$ bln, $z = 50$ friends, $d \approx 5.8$.

NATIONAL RESEARCH
UNIVERSITY