



Social Network Analysis: Data collection III

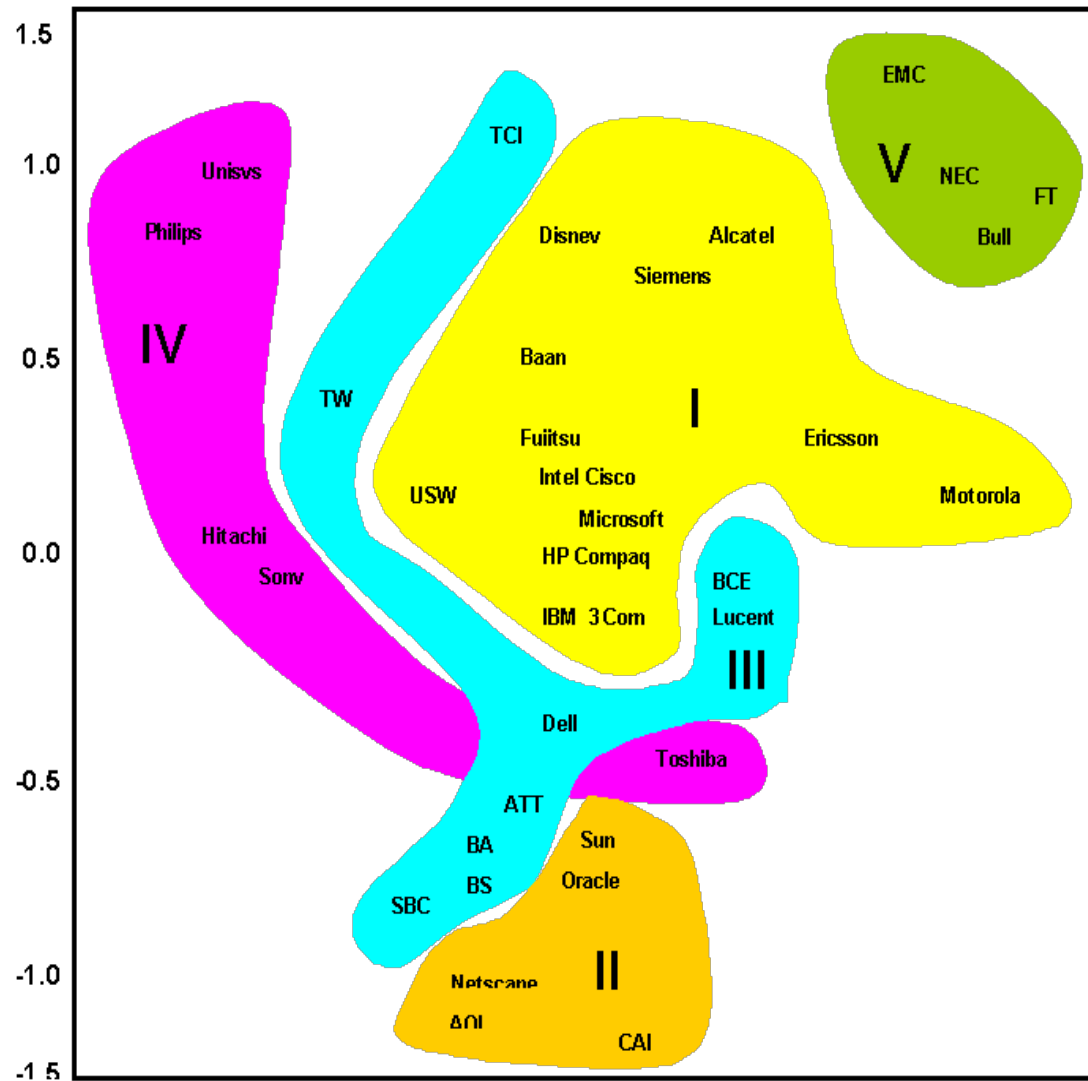
Song Yang, Ph.D.

University of Arkansas

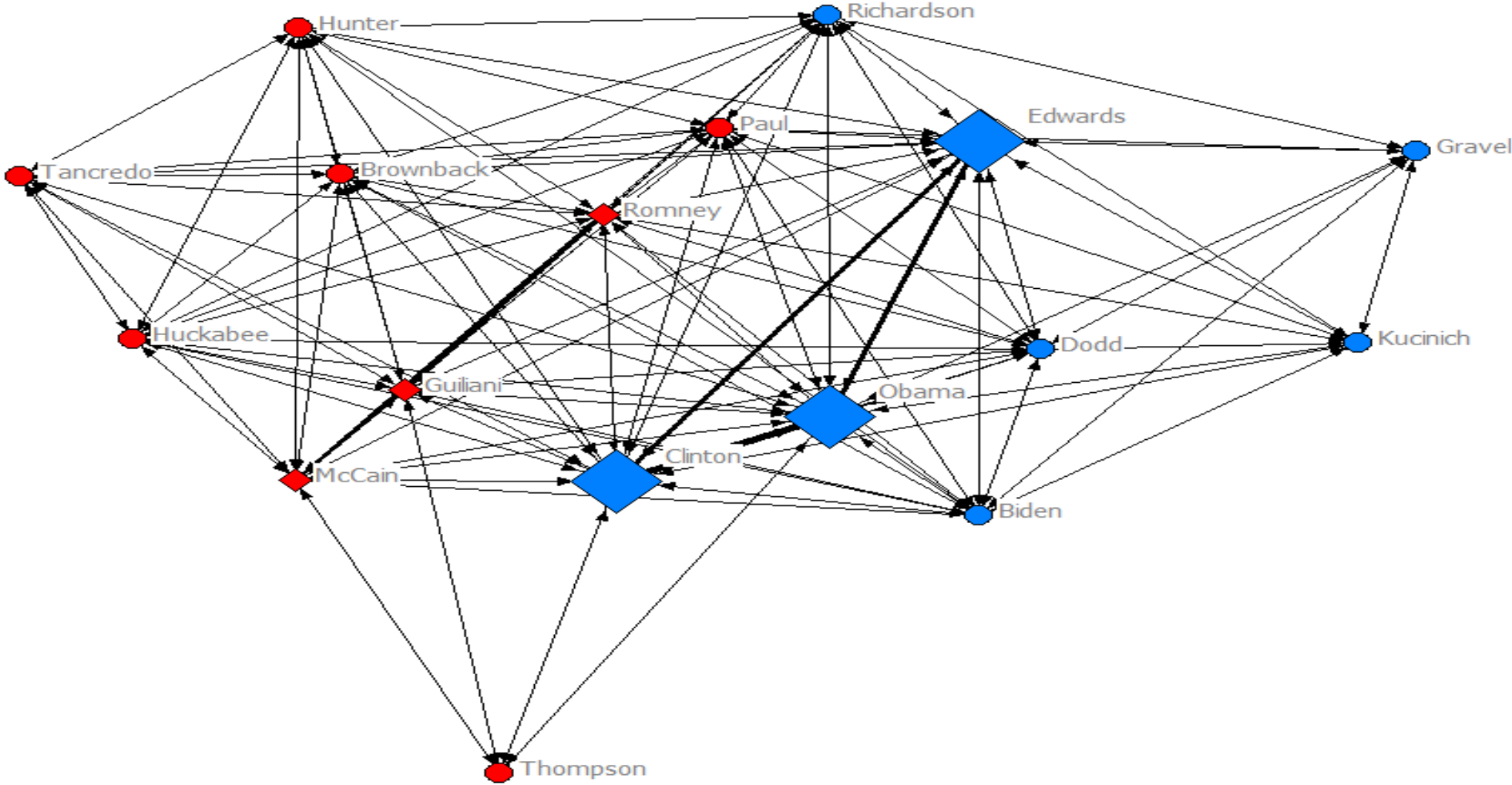
Archival data

- Unearth unobtrusive observational data from already existing information source without posing questions to the actors themselves
- Sources of archival data
 - Personal letters
 - Published books, articles, and patents (citation list)
 - Online data repository: Lexis/Nexis, Free Adger, SDC Platinum, FEC
- Key issue is data mining technique: how to mine massive amount of big data to procure data for researchers' needs: this is where social science and hard sciences (EE or computer science) work together a lot

Strategic alliances in the 1998 core GIS



Preprimary presidential election donor data



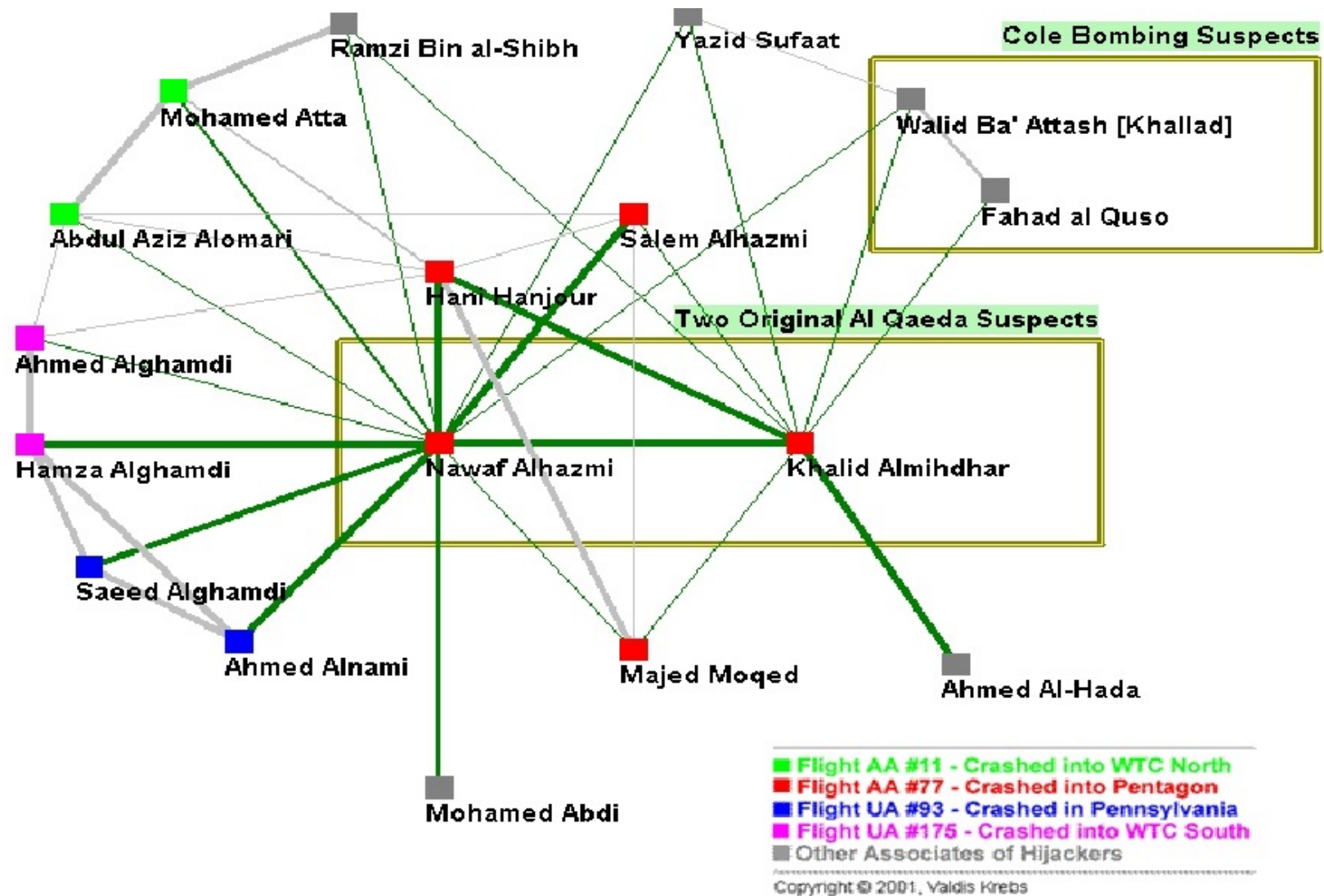
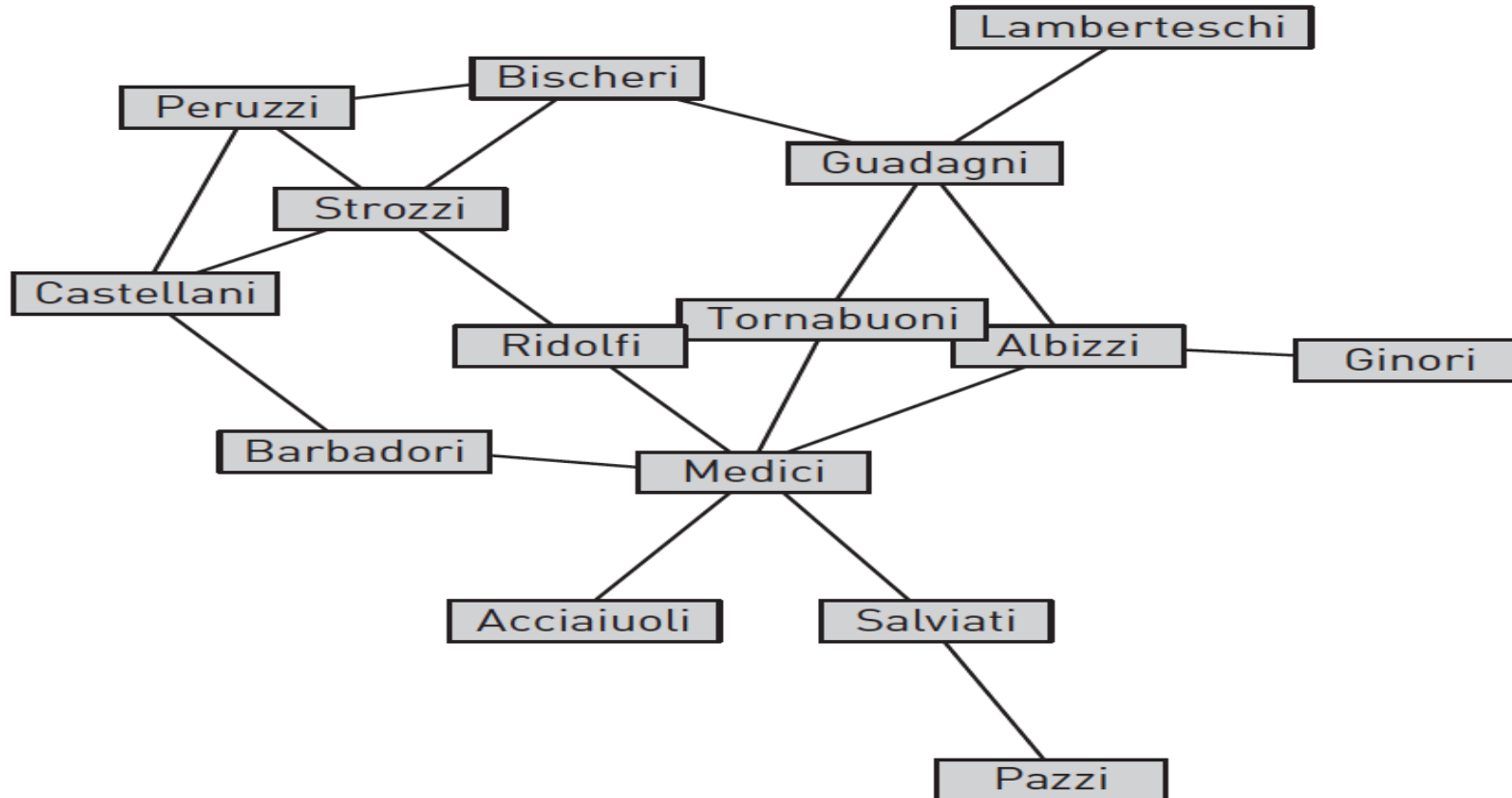


Figure 2 - All nodes within 1 step [direct link] of original suspects

Historians: the Rise of Medici Family

FIGURE 2.3 • **Intermarriage Network of the Ruling Families of Florence**



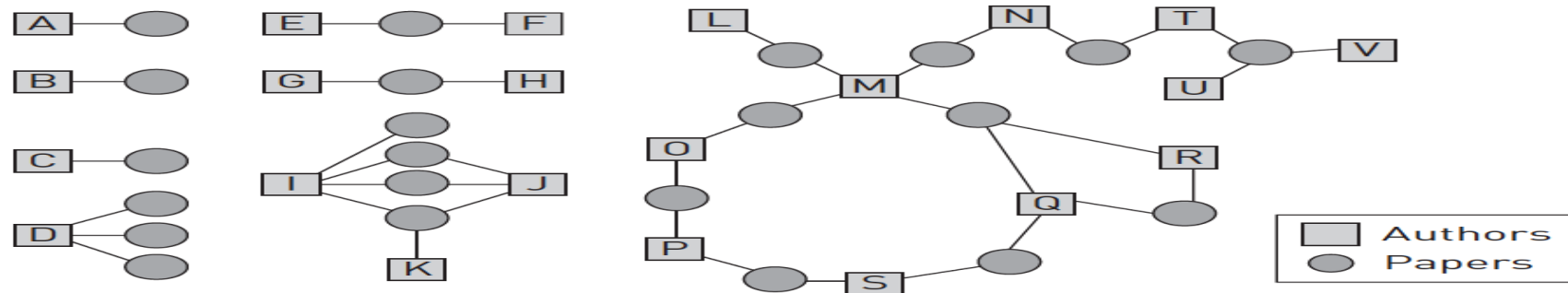
Notes: The nodes represent the 15th century Florentine families, and the ties represent marital relations among those families. Adapted from p. 48 of Newman, M. E. (2010). *Networks: An introduction*. New York, NY: Oxford University Press.

Collaboration and citation networks

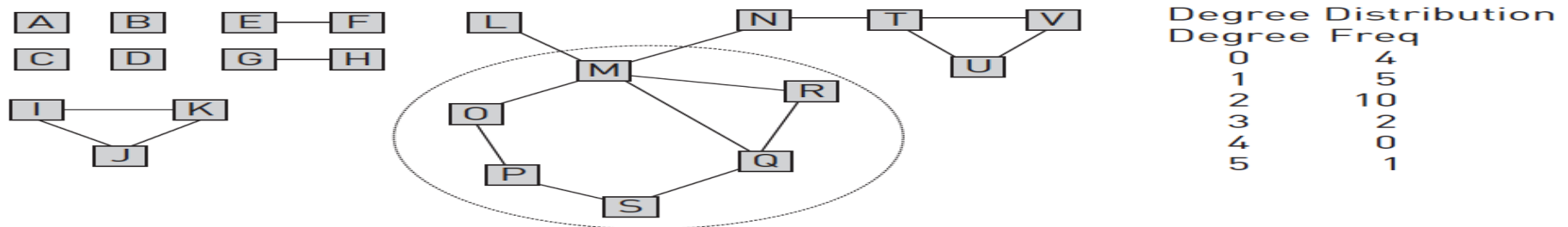
- Individual publications and conversion from bipartite to one-mode network

FIGURE 2.4 • Moody's Construction of Collaboration Networks

a) Individual Publications



b) Collaboration Network



Note: Adapted from p. 220 of Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American Sociological Review*, 69, 213–238.

Citation
network
(whom they
cite and for
what
reasons:
Hargens
2000)

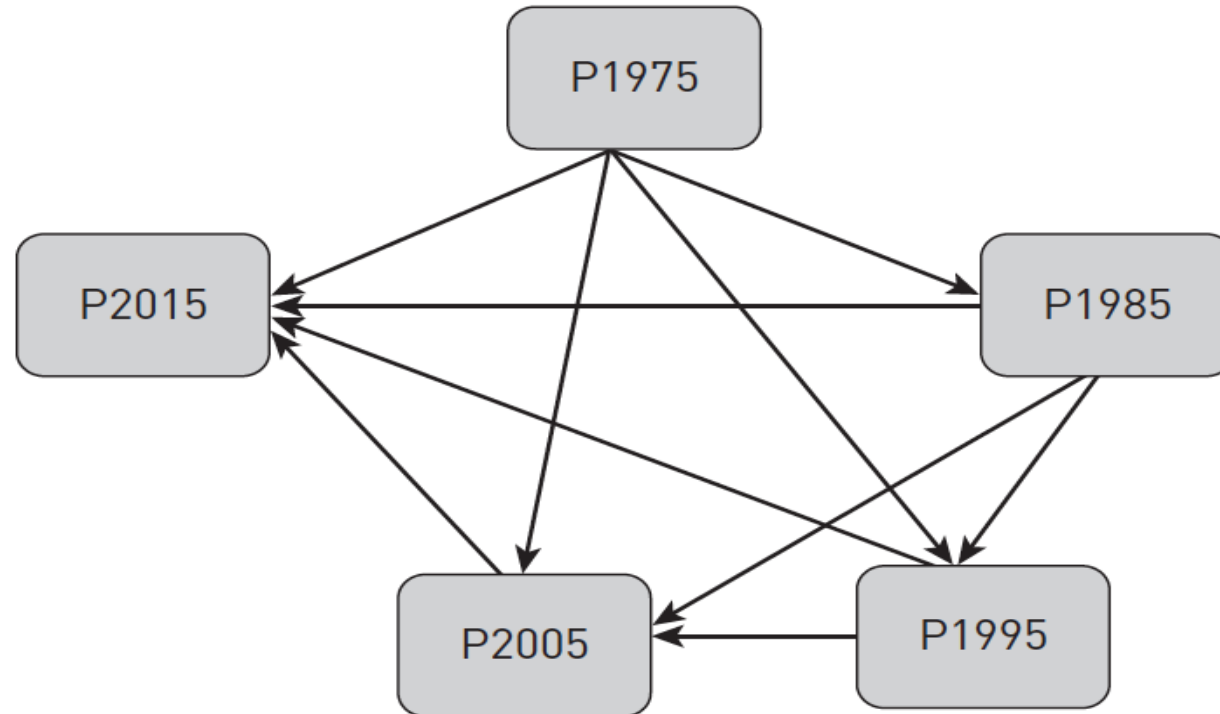
Physical sciences: to reject or correct
specific experimental methods

Social sciences: to align or identify with
or to challenge fundamental
assumptions of the disciplines

Humanities: critiques and
disagreements

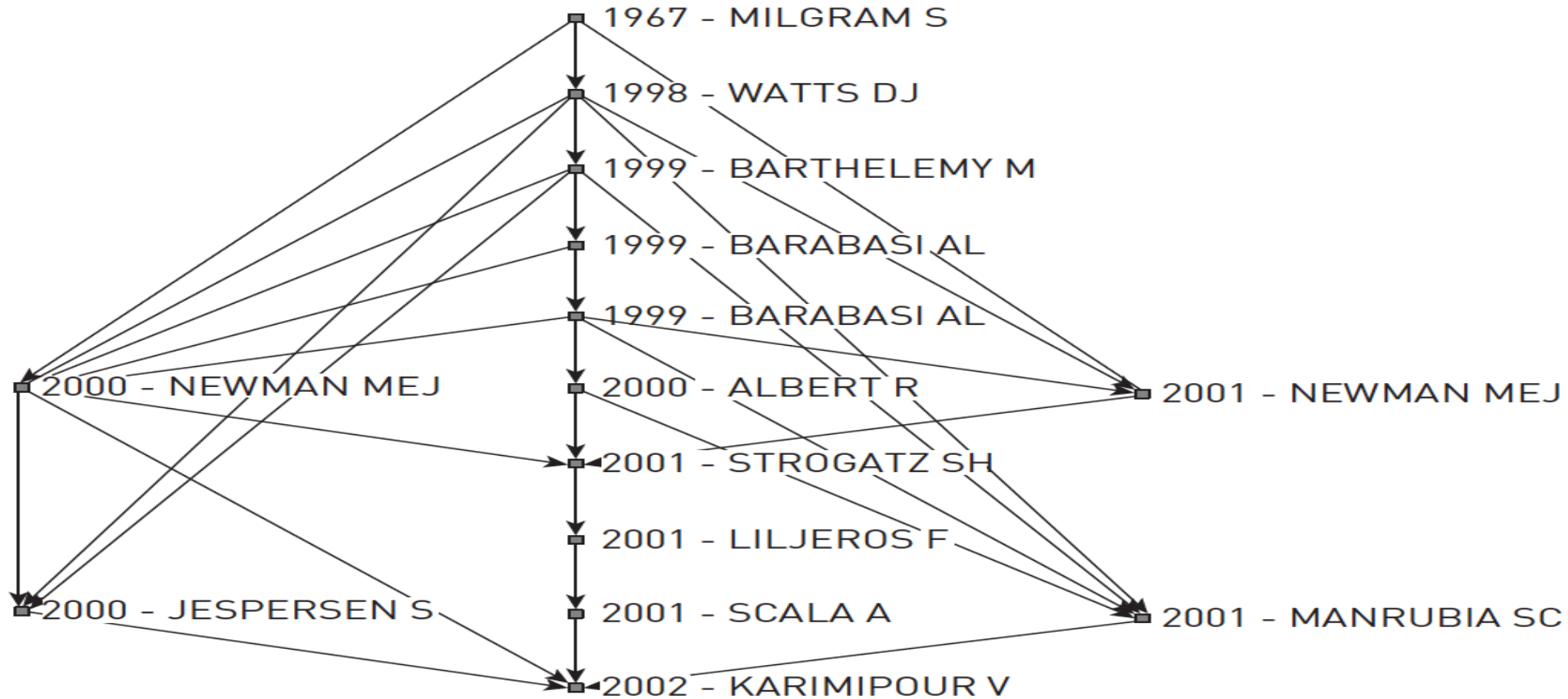
Citation network is acyclical

FIGURE 2.5 • **Graph of Citation Network of Five Papers**



Small world phenomenon

FIGURE 2.6 ● **Historiograph of “Small-World” Papers Descending From Milgram**



Note: Adapted from White, H. D. (2011). Scientific and scholarly networks. In (J. Scott and P. J. Carrington, Eds.), *The SAGE handbook of social network analysis* (pp. 271–286). London, England: Sage Ltd.

Social media and e-communications

Email network (role, content, and time)

Researchers analyzed email exchanges between 158 Enron employees, including its CEOs Kenneth Lay and Jeffrey Skilling

Researchers enhanced the data to include senders and receivers, number of emails exchanged between all pairs, and across different time frames (before, during, and after the crisis)

- During the crisis period, communication among employees became more diverse with respect to established contacts and formal roles. Also
- During the crisis period, previously disconnected employees began to engage in mutual communication, so that interpersonal communication was intensified and spread through the network, bypassing formal chains of communication.

Issues with email social network analysis

Unable to capture the subtleties
that matter a great deal

Many important issues are not
communicated via email exchanges

Other issues

CSSN (computer supported social network)

- Facebook, twitter, Instagram
- WeChat, facetime etc
- Google Scholar
- Dating sites
- Large servers (Stanford and MIT servers)
- Redditor/reddit
- TikTok

Compare and contrast the three types of generators in ego-centric network studies: name generator, positional generator, and resource generator. Could you produce examples in questionnaire surveys that are illustrative of the three generators?

When asked “with whom you play sport?” one respondent produced 20 names in the first interview and 22 names in the second interview. Fifteen names mentioned in the first interview appeared in the second interview. What is the Jaccard coefficient for this respondent?

Questions