

# Small, H. & His Science Mapping



---

Xiaohua Zhou

[Xiaohua.Zhou@drexel.edu](mailto:Xiaohua.Zhou@drexel.edu)

March 16, 2004



# Brief of Small

---

- Joint Ph.D. in Chemistry & History of Science from University of Wisconsin
- A Brief Career as A Historian of Science at the American Institute of Physics' Center
- Join ISI, Philadelphia in 1972
- Propose the Method of Document Co-citation Analysis in 1973
- Published around 25 Papers in Co-citation Analysis and Sciences Mapping since 1973
- JASIS Editorial Board and Fellow of AAAS since 1985
- Award of Merit from ASIST in 1998
- Chief Scientist at ISI





# Selected Document Set

---

- The Geography of Science: Disciplinary and National Mappings (1985)
- Macro-level Changes in the Structure of Co-citation Clusters (1993)
- A General Framework for Creating Large-scale Maps of Science in Two or Three Dimensions: The SciViz System (1998)
- A Passage Through Science: Crossing Disciplinary Boundaries (1999a)
- Visualizing Science by Mapping (1999b)



# Sample Highly Cited Doc

---

- Obtain a fairly representative highly cited documents sample for stable analysis result
- Use fractional citation counts rather than integer count to remove the effect of uneven citation

e.g. for a paper with ten references, each ref. has 0.1 unit of strength.

- Set integer citation threshold to remove random noise

e.g. for a paper cited by 3 docs all with 2 references only, its fractional count is 1.5, while the count of a paper cited by 10 docs all with 10 references is 1.0 only.

# Doc Co-citation Clustering

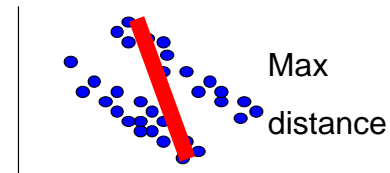
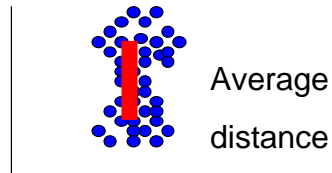
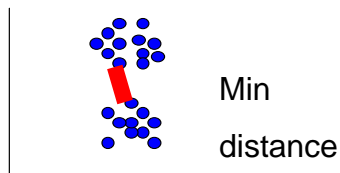
- Normalized co-citation frequency as linkage strength

$$Freq = AB / \sqrt{A * B}$$

- Four iterations hierarchical clustering C1-C5
- Single-linkage algorithm

Simplicity of implementation for massive files  
loose connected, highly chained

Cut off cluster size, variable similarity threshold  
Isolated documents or clusters





# Ordination & Display

---

- Multidimensional scaling for ordination
- Triangulation for ordination

fast, less computational intensive  
depending on the order the object assigned to the map

- Common coordinate space

bind individual clusters into a global common coordinate space  
without overlap

- Volvox display

smaller, lower-level objects are represented as circles within  
larger, higher-level objects



# Macro Evolution of Sciences

---

- Method & Result

use cluster strings to correspond between clusters across years  
competitive bind among fields, analogous to atoms competing for a binding site on a molecule

discover a “pulsating model” of knowledge evolution

- Remaining Work

Use other co-citation method to verify the finding

Intuitively display the evolution of science

Extend to micro-level evolution of sciences

Improve or replace single-linkage clustering



# Information Pathway

---

- An automated method for creating information pathways with arbitrarily beginning and ending point.
- Interpretation of links in the pathway

Content analysis of co-citation passages for pairs of document along the information path.

A close examination of titles of the linked documents

- Implications for information retrieval, science unity, discovery, epistemology, and evaluation



# Contributions

---

- A framework and an implementation of large-scale document co-citation based science mapping
- Longitudinal co-citation analysis
- Information Pathway Discovery
- Implications for bibliometrics and information visualization, for scientific policy makers, researchers and profession of information sciences



# Remaining Work

---

- Improve or replace single linkage clustering method
- Better display

enhance volvox display

intuitively represent the evolution of sciences

visualize the information pathways across hierarchical levels

- The work on longitudinal co-citation analysis and information pathway is still at infant stage