

# Analysis of InfoVis 2004 Contest Data: A Survey & Analysis

Chia-Ning Chiang  
Ron Jung-Rung Han  
April 21, 2004

April 21, 2004

1

## Tasks: Survey and Analysis

- This is a survey project: to evaluate existing visualization tools, in particular, VxInsight, on their capabilities to explore the dataset and perform the InfoVis contest tasks
- We use several light-weight tools to explore the possibilities of each of the InfoVis 2004 Contest task questions.
- We focus on VxInsight, a commercial grade visualization tool developed by the Sandia National Laboratories, to do these contest tasks

April 21, 2004

2

## Related Work

- Visualizing Science by Citation Mapping
  - Citation analysis, term co-occurrence, co-classification, or Author co-citation analysis, etc.
  - Traditionally, standard output for literature mapping studies has been a circle plot
- Visualization Tools
  - VR-VIBE
  - Cat-a-Cone
  - SCI-Map
  - SPIRE
  - VxInsight

April 21, 2004

3

## Exploration of Tools Before Using VxInsight

- Bibexcel (<http://www.umu.se/inforsk/Bibexcel/>)
  - This tool-box is designed to assist a user in analyzing bibliographic data.
  - It generates the counts and coordination required to generate the citation map, a Multidimensional Scaling map.
  - This helps us understand the data relationships.
- AuthorLink (<http://cite.cis.drexel.edu>)
  - It creates live interfaces for cocited author retrieval on the fly.
  - This help us to appreciate what a well-structured visualization of author co-citation maps can be like.

April 21, 2004

4

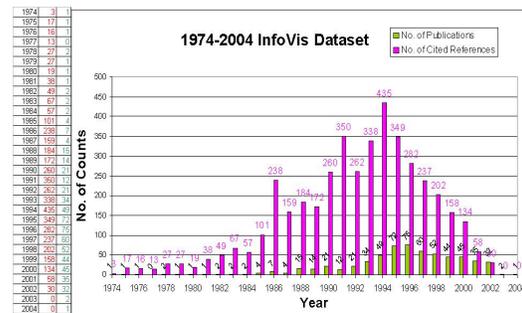
## Use of VxInsight

- Parse XML data
- Set up VxInsight
  - Provide several files required by VxInsight. The most noteworthy one is the input dataset that is put in an ODBC-compliant database.
  - Specify a similarity measure between data objects.
  - Generate [x,y] positions for each data object from the similarity measure. This step, known as ordination, is done by VxOrd (a subsystem of VxInsight). A similarity measure is optional.
  - Load the data into VxInsight.

April 21, 2004

5

## An Overview of the Dataset



April 21, 2004

6

## Exploration of the Datasets (1 of 2)

- The Research topic trends from 1974 to 2004
  - Interface/approach→structure/approach→visualization/information→visualization/data
  - The most cited is Cone Tree, Fiseye Views second, followed by The Perspective Walls
  - IEEE produced more papers than ACM
  - In top 10 most cited papers, 8 from ACM and 1 from IEEE
  - Tufte is the only one book in the top 10
- The Author co-citation analysis and the Citation Context from ISI's Web of Science and CiteSeer
  - The most cited articles comes from Xerox Palo Alto Research Center and University of Maryland.

April 21, 2004

7

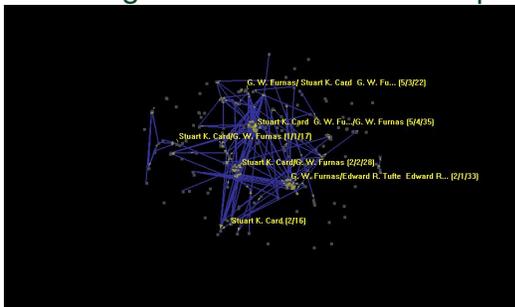
## Exploration of the Datasets (2 of 2)

	InfoVis	ISI Web of Science	CiteSeer
Stuart Card	65	68	64
G. W. Furnas	61	171	237
Ahlberg & Shneiderman	24	76	153
Johnson & Shneiderman	24	9	0

April 21, 2004

8

## VxInsight Author-cocitation Map



April 21, 2004

9

## Pathfinder Networks Author Co-citation Map

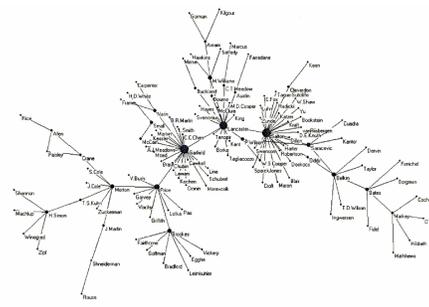


FIG. 1. PFNET of 121 information science authors based on raw cocitation counts.

April 21, 2004

10

## User Study

- 15 subjects in library, archival and information science.
- Incredible, wonderful, but need explanation of the metaphor
- 2 subjects did a hands-on tests, very intuitive, but not all the functions. (even ourselves, the demonstrators)
- Galaxy view's color coding should be consistent with the terrain view.

April 21, 2004

11

## Lessons Learned

- The Dataset
- The Software Tools
- The Process
- Cooperation between two different mind-sets
- Good support from Brian Wylie

April 21, 2004

12