Visualization of CRAN Package Dependency Network

Group 3: X Zhang, X Chen, W Qiwen, N Xiao

2010 PKU Vis Summer School

August 18, 2010
Slides now available at:

http://vis.pku.edu.cn/wiki/

Public Courses ⇒ Summer School s10 ⇒ Group 3
Outline

1. Why
2. How
3. What we got
4. Reference
5. Appendix
What is CRAN?
Visualization of CRAN Package Dependency Network

Why

Comprehensive R Archive Network
CTAN: Comprehensive TeX Archive Network

CPAN: Comprehensive Perl Archive Network
Why needs to be VISUALIZED?
Visualization of CRAN Package Dependency Network

Why

1. Real world data beside us
2. Strengthen the open source community
3. Has never been **VISUALIZED** before
Data Extracted from CRAN: Raw Text.

After Pre-Processing: 1805 Nodes, 4020 Edges.
Now available at Vis wiki.
Visualization of CRAN Package Dependency Network

How

Tools:
1. Pajek (1996)
2. Gephi (2008)
3. Network Manipulator (McGuffin, 2009)

Why not R
What we got

1. Basic SNA
2. “Big Picture”
3. Manual Analysis
Visualization of CRAN Package Dependency Network

What we got

Basic SNA

Figure: Output Degree of the Network
Basic SNA

Diameter:
\texttt{stats} \Rightarrow \texttt{tossm}

Distance: 5.
Clustering Coefficients:

$$CC(i) = \frac{2E_1(i)}{(k_i(k_i - 1))}$$

$$CC = \sum_{i=1}^{N} \frac{CC(i)}{N}$$

For this network: $CC = 0.0697$
Visualization of CRAN Package Dependency Network

Hubs and Authority

Figure: Hubs
Hubs and Authority

Figure: Authority
“Big Picture”

Gephi project and PDF(Vector) Version also available at Vis wiki.
Force Directed Layout

Figure: Force Directed Graph by GUESS
Manual Analysis: Most Important Packages

Figure: Manual Analysis: Most Important Packages (1)
Figure: Manual Analysis: Most Important Packages (2)
Figure: Manual Analysis: Most Important Packages (3)
Figure: Manual Analysis: Most Important Packages (4)
Figure: Manual Analysis: Most Important Packages (5)
Figure: Manual Analysis: Most Important Packages (6)
Visualization of CRAN Package Dependency Network

What we got

Figure: Manual Analysis: Most Important Packages (7)
Visualization of CRAN Package Dependency Network

What we got

Figure: Clustering Result by MCL

Group 3: X Zhang, X Chen, W Qiwen, N Xiao
Future Jobs

1-mode $\Rightarrow$ 2-mode

Authors + Packages
The Namecards

Zhang Xin
Visualization of CRAN Package Dependency Network

Appendix

The Namecards

Chen Xuan

Group 3: X Zhang, X Chen, W Qiwen, N Xiao
The Namecards

Qiwen Wang

Group 3: X Zhang, X Chen, W Qiwen, N Xiao
The Namecards

Visualization of CRAN Package Dependency Network

Group 3: X Zhang, X Chen, W Qiwen, N Xiao
Questions?
Thanks