

Social Computing and Big Data Analytics

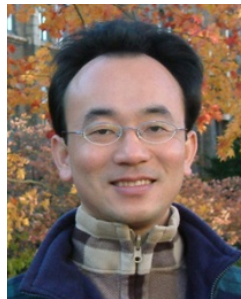
社群運算與大數據分析

Social Network Analysis (社會網絡分析)

1042SCBDA12

MIS MBA (M2226) (8628)

Wed, 8,9, (15:10-17:00) (B309)



Min-Yuh Day

戴敏育

Assistant Professor

專任助理教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2016-05-18



課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
1	2016/02/17	Course Orientation for Social Computing and Big Data Analytics (社群運算與大數據分析課程介紹)
2	2016/02/24	Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data (資料科學與大數據分析： 探索、分析、視覺化與呈現資料)
3	2016/03/02	Fundamental Big Data: MapReduce Paradigm, Hadoop and Spark Ecosystem (大數據基礎：MapReduce典範、 Hadoop與Spark生態系統)

課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
4	2016/03/09	Big Data Processing Platforms with SMACK: Spark, Mesos, Akka, Cassandra and Kafka (大數據處理平台SMACK : Spark, Mesos, Akka, Cassandra, Kafka)
5	2016/03/16	Big Data Analytics with Numpy in Python (Python Numpy 大數據分析)
6	2016/03/23	Finance Big Data Analytics with Pandas in Python (Python Pandas 財務大數據分析)
7	2016/03/30	Text Mining Techniques and Natural Language Processing (文字探勘分析技術與自然語言處理)
8	2016/04/06	Off-campus study (教學行政觀摩日)

課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
9	2016/04/13	Social Media Marketing Analytics (社群媒體行銷分析)
10	2016/04/20	期中報告 (Midterm Project Report)
11	2016/04/27	Deep Learning with Theano and Keras in Python (Python Theano 和 Keras 深度學習)
12	2016/05/04	Deep Learning with Google TensorFlow (Google TensorFlow 深度學習)
13	2016/05/11	Sentiment Analysis on Social Media with Deep Learning (深度學習社群媒體情感分析)

課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
14	2016/05/18	Social Network Analysis (社會網絡分析)
15	2016/05/25	Measurements of Social Network (社會網絡量測)
16	2016/06/01	Tools of Social Network Analysis (社會網絡分析工具)
17	2016/06/08	Final Project Presentation I (期末報告 I)
18	2016/06/15	Final Project Presentation II (期末報告 II)

Social Computing

Social Network Analysis

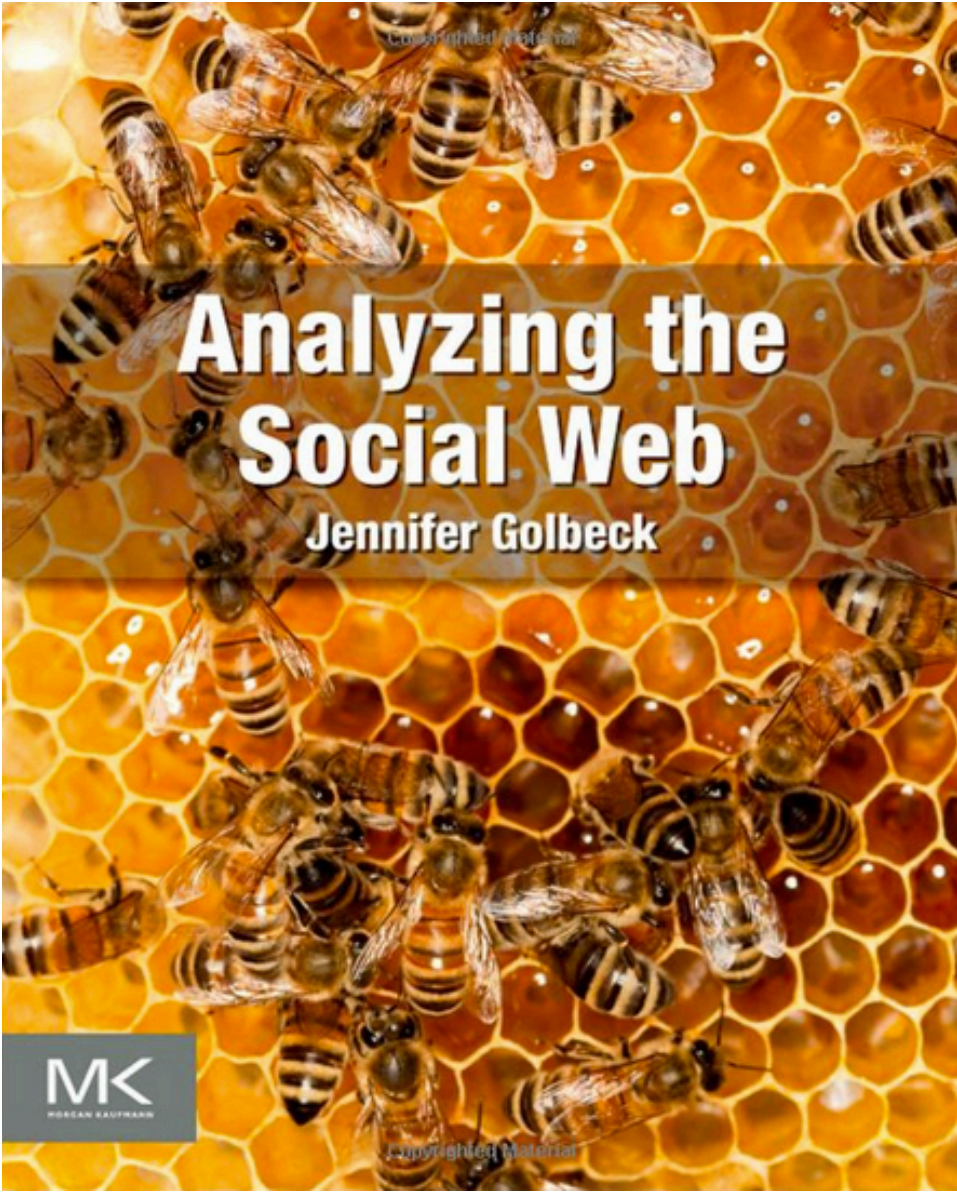
Social Computing

- Social Network Analysis
- Link mining
- Community Detection
- Social Recommendation

Business Insights
with
Social Analytics

Analyzing the Social Web: Social Network Analysis

Jennifer Golbeck (2013), *Analyzing the Social Web*, Morgan Kaufmann



Social Network Analysis (SNA)

Facebook TouchGraph

TouchGraph Photos x

box.touchgraph.com/facebook/TGFacebookBrowser.php?&signed_request=Gi-L3_6HrZ0S3SjxAXGdHR0rhMzqBjUnvFJ9vE4W6vg.eyJhbGdvcm0aG0iOiJITUFDI☆

Profiles Networks

Show Top 100 Friends Show All Friends Upload Advanced Restart

Zoom: Spacing:

Min-Yuh Day
 Networks: None
 Mutual Friends: 681

Facebook Profile

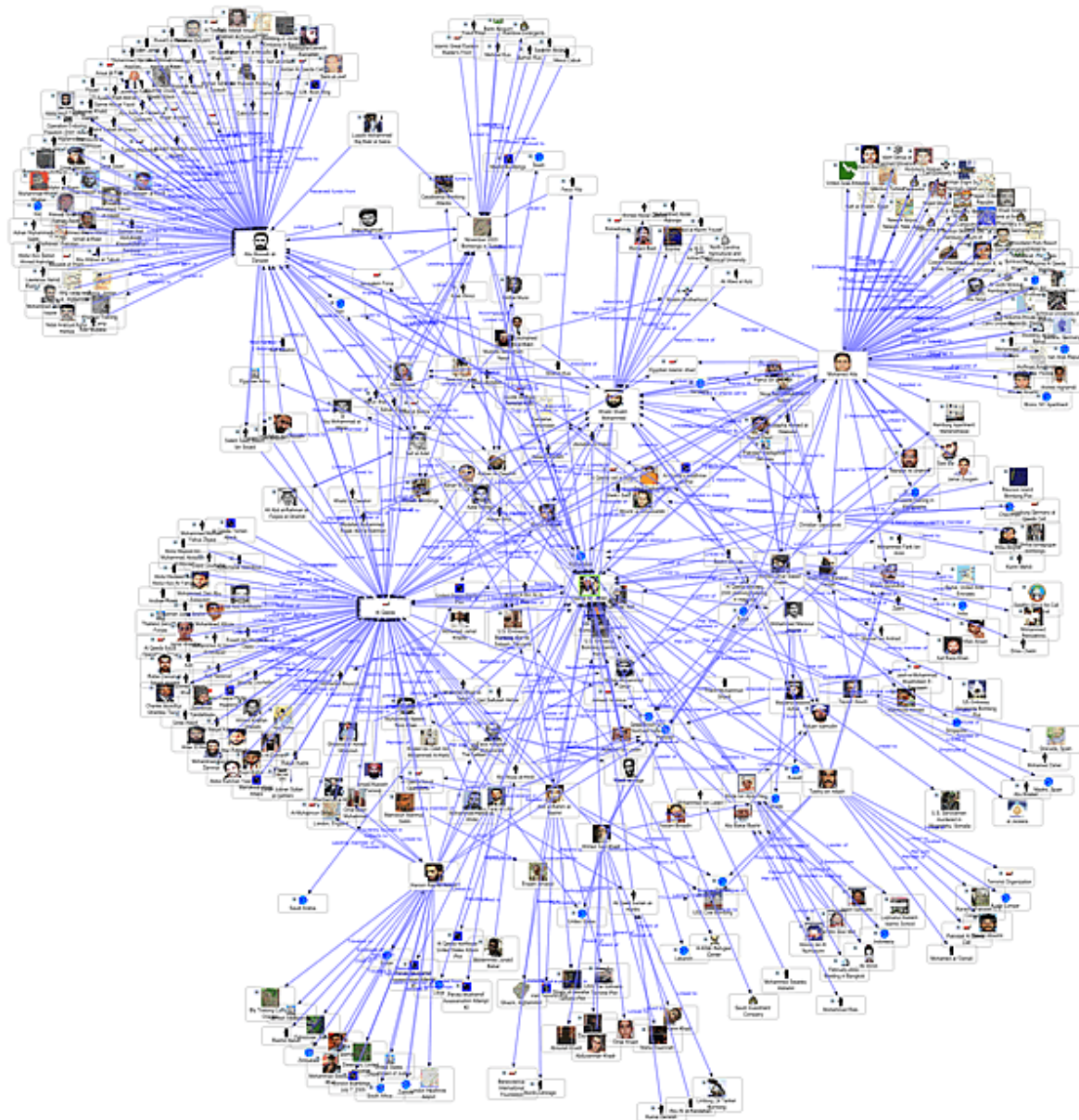
Network All All List Photo

Name	Rank #	Friend #	民
Min-Yuh Day	1	681	
Gladys Hsieh	2	85	
黃西田	3	74	
施盛賓	4	67	
John Lee	5	104	
Kevin Tu	6	61	
Yung Yu Shih	7	45	
Wei Chen	8	107	
Chichang Jou	9	50	
Allen Green	10	81	
黃煒勳	11	65	
梁德昭	12	44	
Eric Chen	13	51	
吳錦波	14	39	
Jessica Tien	15	49	
蔡名宜	16	112	
Enrico Lu	17	59	
YaHan Hsieh	18	64	
王慧雯	19	56	
薛聖譚	20	80	
蝦米	21	73	

ICCU

powered by TouchGraph

Social Network Analysis



Social Network Analysis

- A **social network** is a social structure of people, related (**directly** or **indirectly**) to each other through a common relation or interest
- **Social network analysis (SNA)** is the study of social networks to understand their **structure** and **behavior**

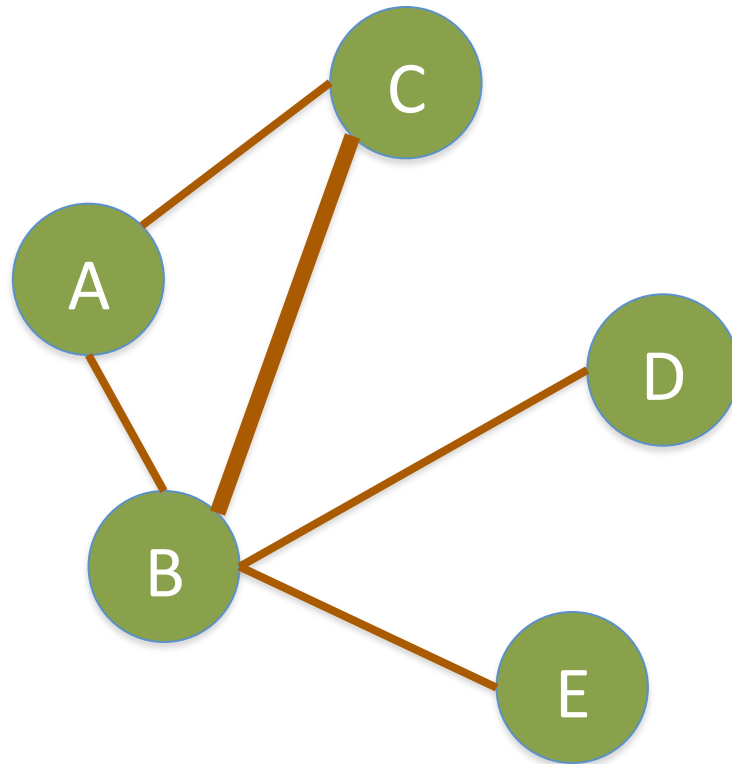
Social Network Analysis (SNA)

Centrality

Prestige

Graph Theory

Graph



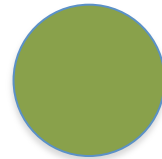
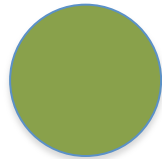
Graph

$$g = (V, E)$$

Vertex (Node)



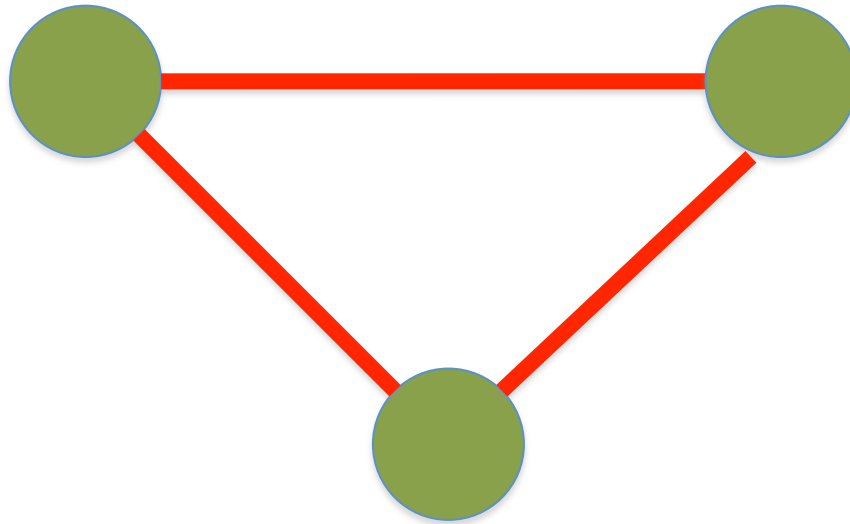
Vertices (Nodes)



Edge



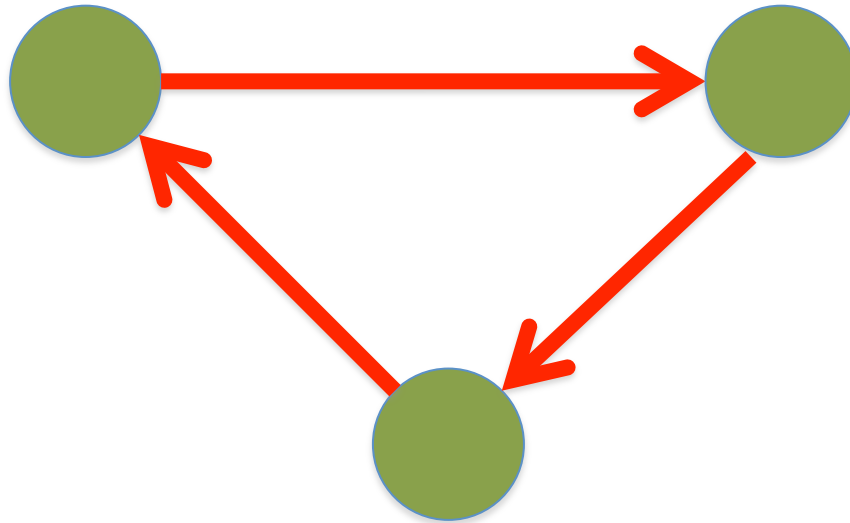
Edges



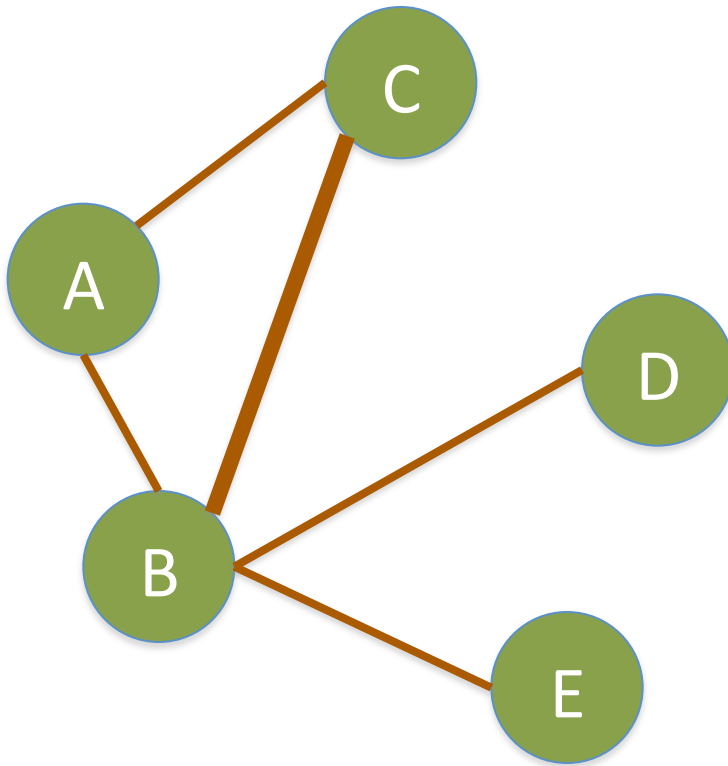
Arc



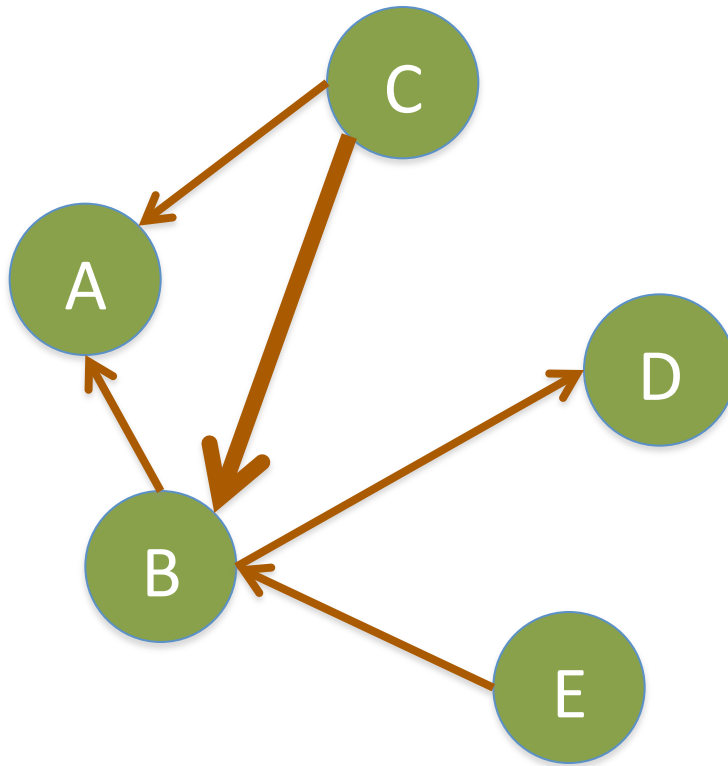
Arcs



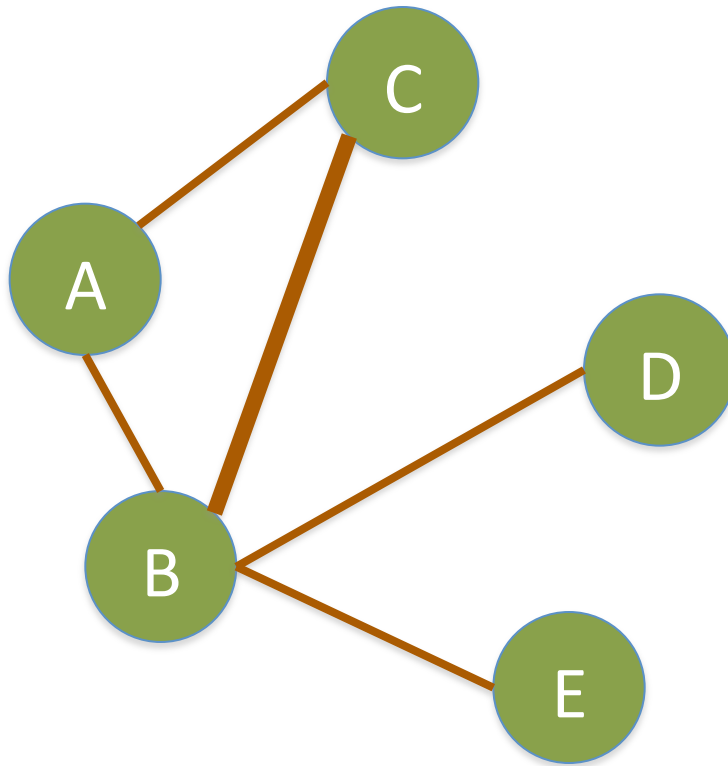
Undirected Graph



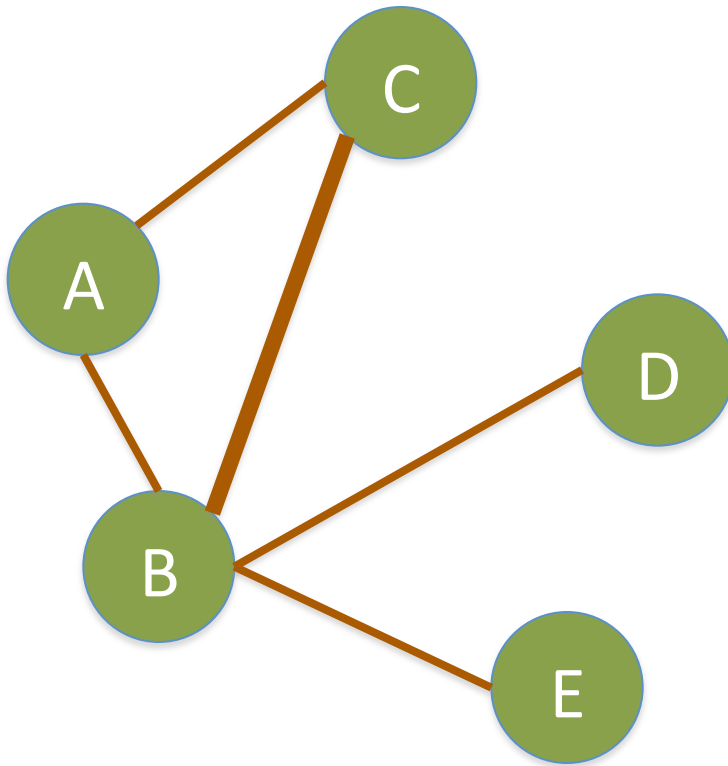
Directed Graph



Degree

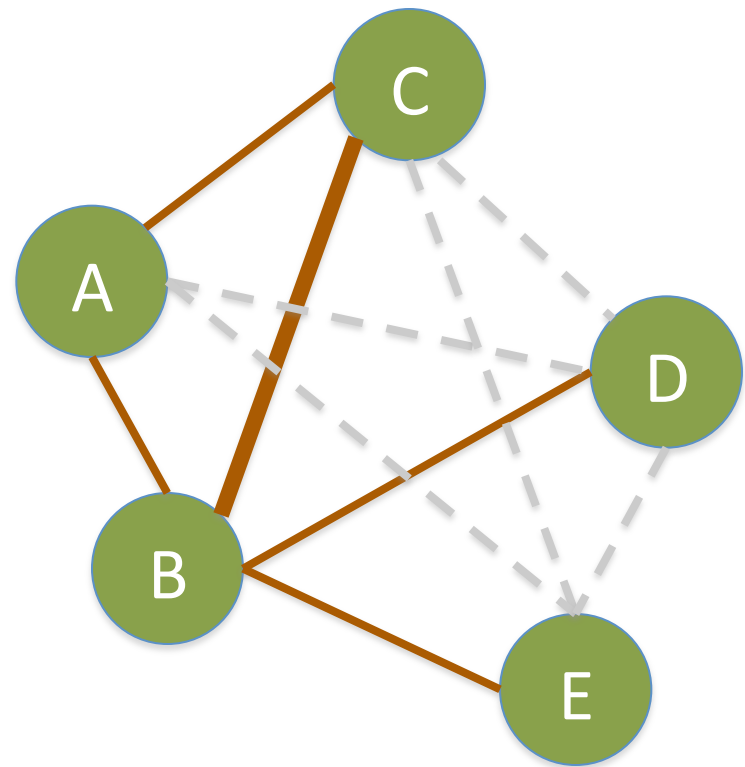


Degree



A: 2
B: 4
C: 2
D: 1
E: 1

Density

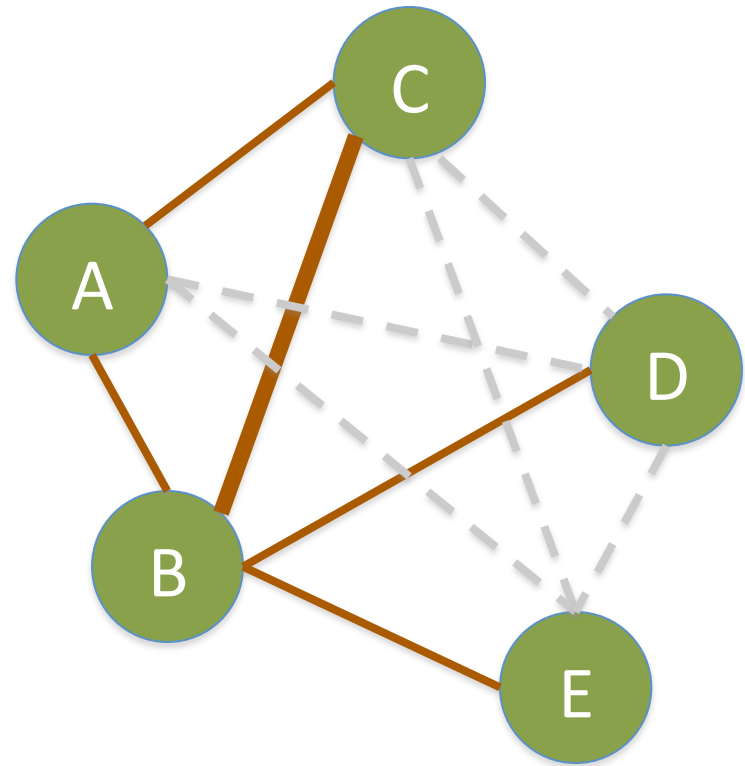


Density

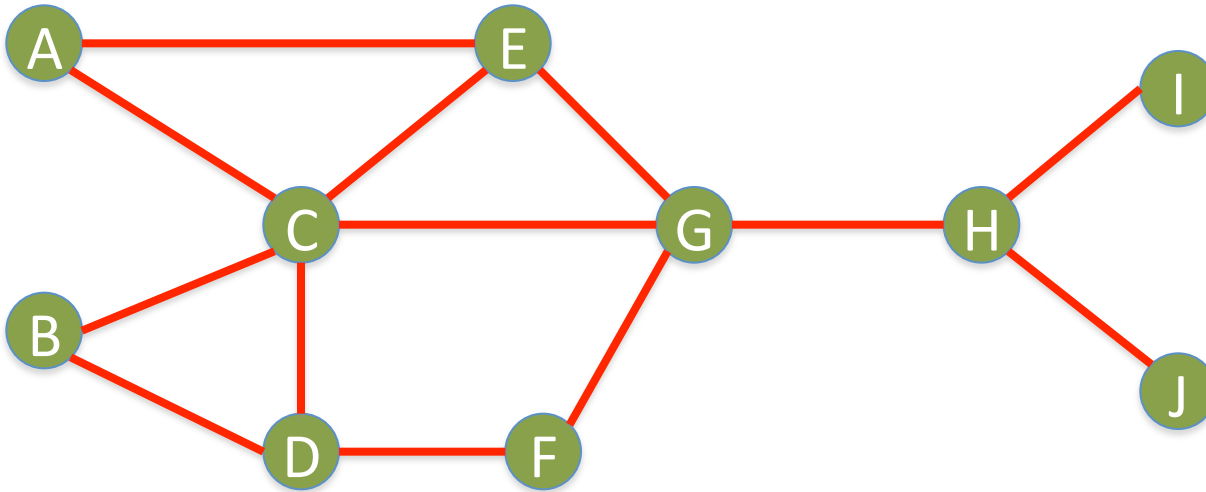
Edges (Links): 5

Total Possible Edges: 10

Density: $5/10 = 0.5$



Density



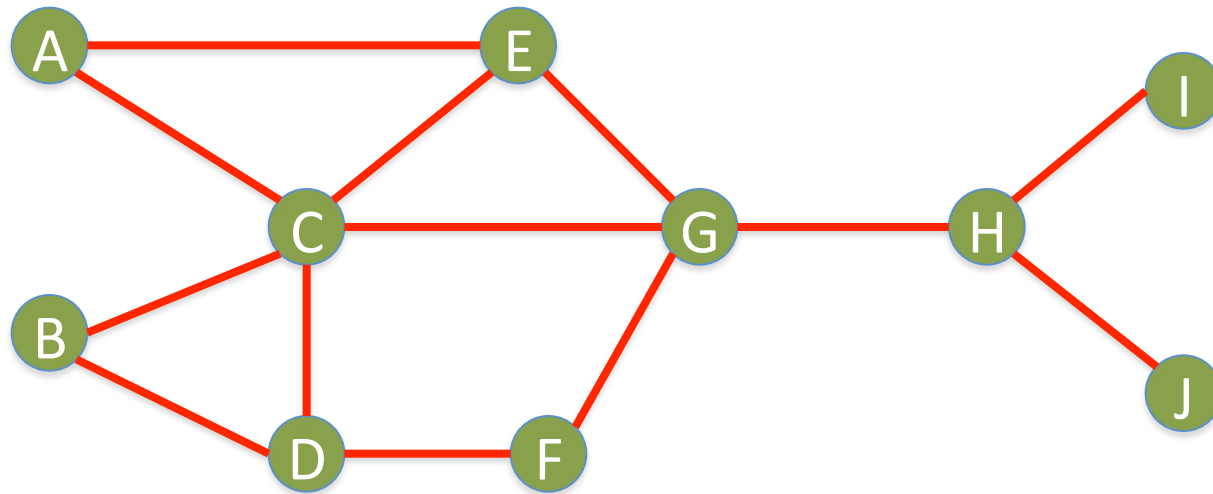
Nodes (n): 10

Edges (Links): 13

Total Possible Edges: $(n * (n-1)) / 2 = (10 * 9) / 2 = 45$

Density: $13/45 = 0.29$

Which Node is Most **Important**?



Centrality

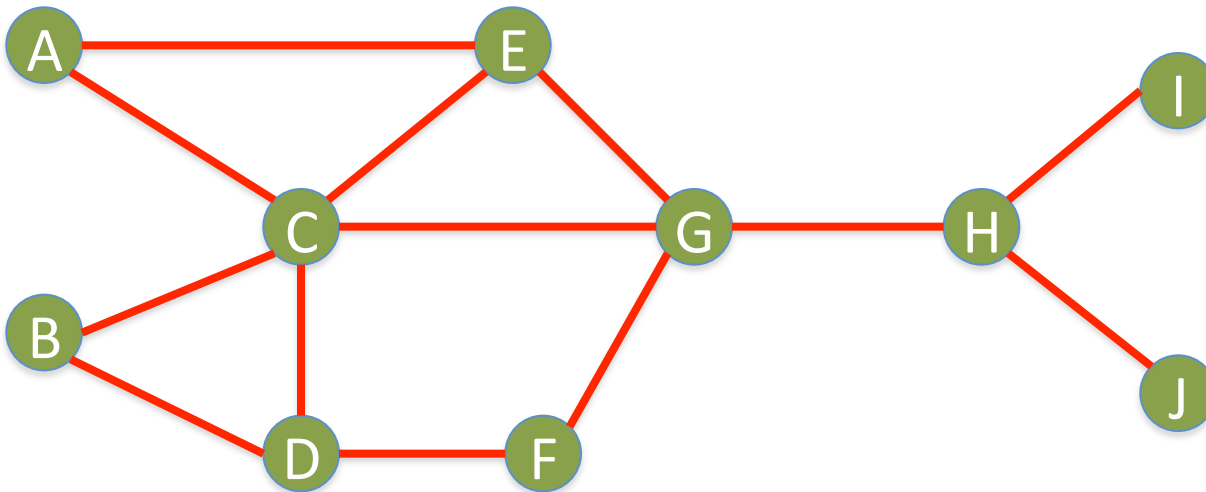
- **Important or prominent actors** are those that are linked or involved with other actors extensively.
- A person with extensive contacts (links) or communications with many other people in the organization is considered more important than a person with relatively fewer contacts.
- The links can also be called **ties**.
A **central actor** is one involved in many ties.

Social Network Analysis (SNA)

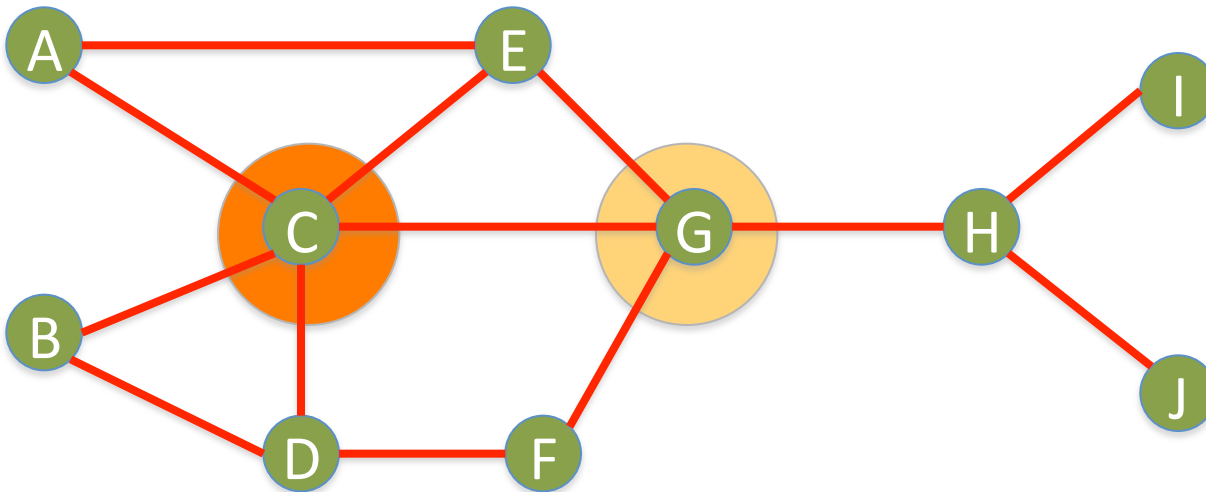
- Degree Centrality
- Betweenness Centrality
- Closeness Centrality

Degree Centrality

Social Network Analysis: Degree Centrality



Social Network Analysis: Degree Centrality



Node	Score	Standardized Score
A	2	$2/10 = 0.2$
B	2	$2/10 = 0.2$
C	5	$5/10 = 0.5$
D	3	$3/10 = 0.3$
E	3	$3/10 = 0.3$
F	2	$2/10 = 0.2$
G	4	$4/10 = 0.4$
H	3	$3/10 = 0.3$
I	1	$1/10 = 0.1$
J	1	$1/10 = 0.1$

Betweenness Centrality

Betweenness centrality:

Connectivity

Number of shortest paths
going through the actor

Betweenness Centrality

$$C_B(i) = \sum_{j < k} g_{jk}(i) / g_{jk}$$

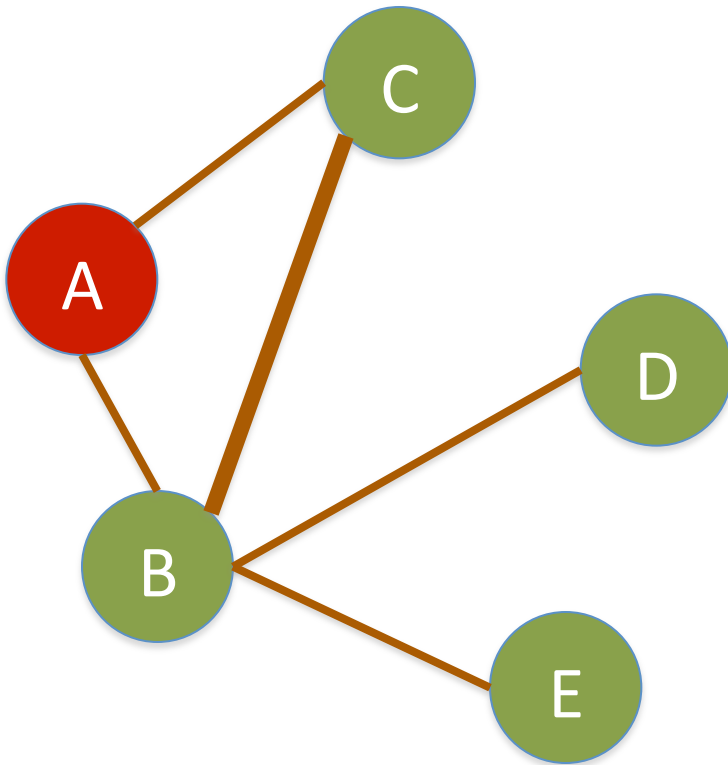
Where g_{jk} = the number of shortest paths connecting jk
 $g_{jk}(i)$ = the number that actor i is on.

Normalized Betweenness Centrality

$$C'_B(i) = C_B(i) / [(n-1)(n-2) / 2]$$

**Number of pairs of vertices
excluding the vertex itself**

Betweenness Centrality



A:

$$B \rightarrow C: 0/1 = 0$$

$$B \rightarrow D: 0/1 = 0$$

$$B \rightarrow E: 0/1 = 0$$

$$C \rightarrow D: 0/1 = 0$$

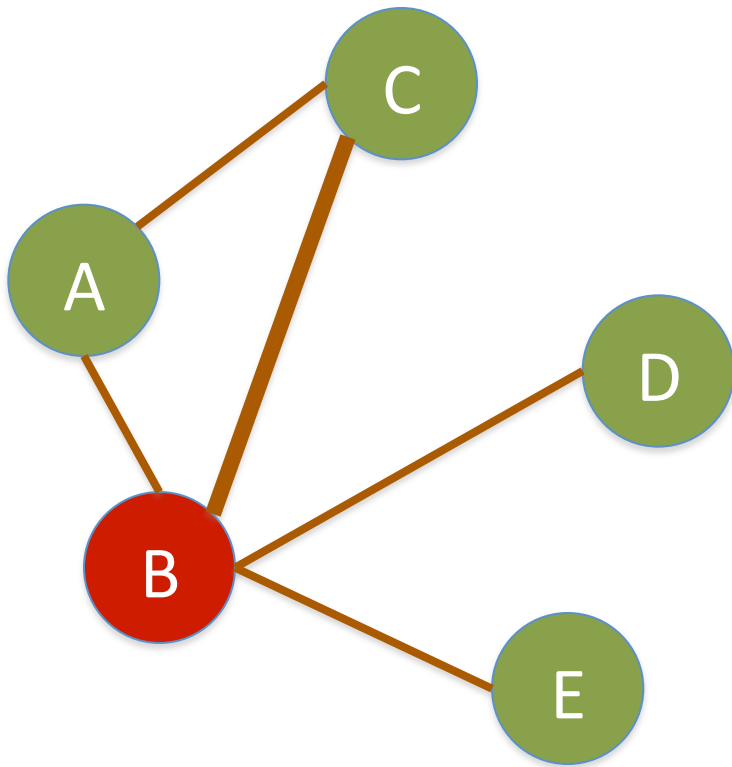
$$C \rightarrow E: 0/1 = 0$$

$$D \rightarrow E: 0/1 = 0$$

Total: 0

A: Betweenness Centrality = 0

Betweenness Centrality



B:

$$A \rightarrow C: 0/1 = 0$$

$$A \rightarrow D: 1/1 = 1$$

$$A \rightarrow E: 1/1 = 1$$

$$C \rightarrow D: 1/1 = 1$$

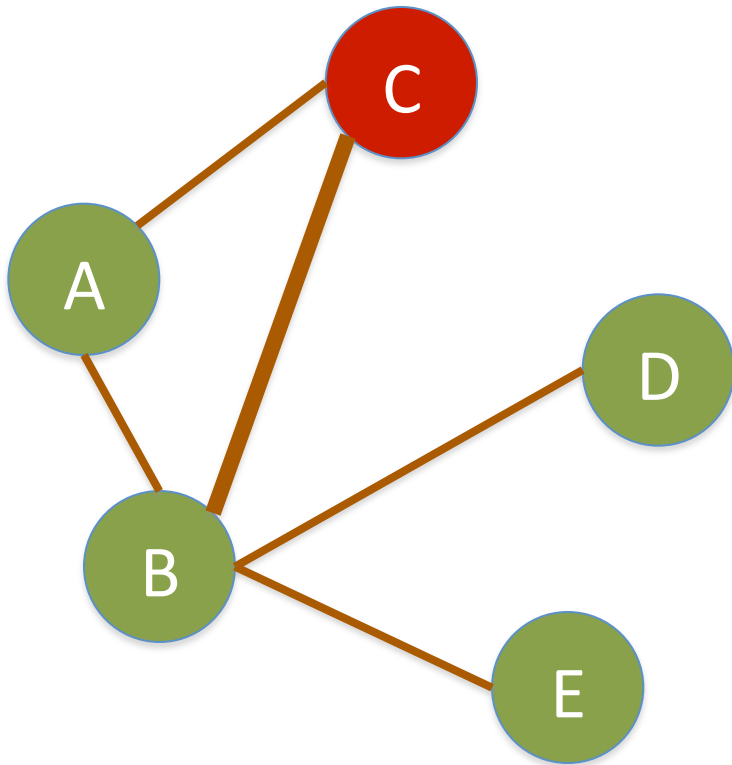
$$C \rightarrow E: 1/1 = 1$$

$$D \rightarrow E: 1/1 = 1$$

Total: 5

B: Betweenness Centrality = 5

Betweenness Centrality



C:

$$A \rightarrow B: 0/1 = 0$$

$$A \rightarrow D: 0/1 = 0$$

$$A \rightarrow E: 0/1 = 0$$

$$B \rightarrow D: 0/1 = 0$$

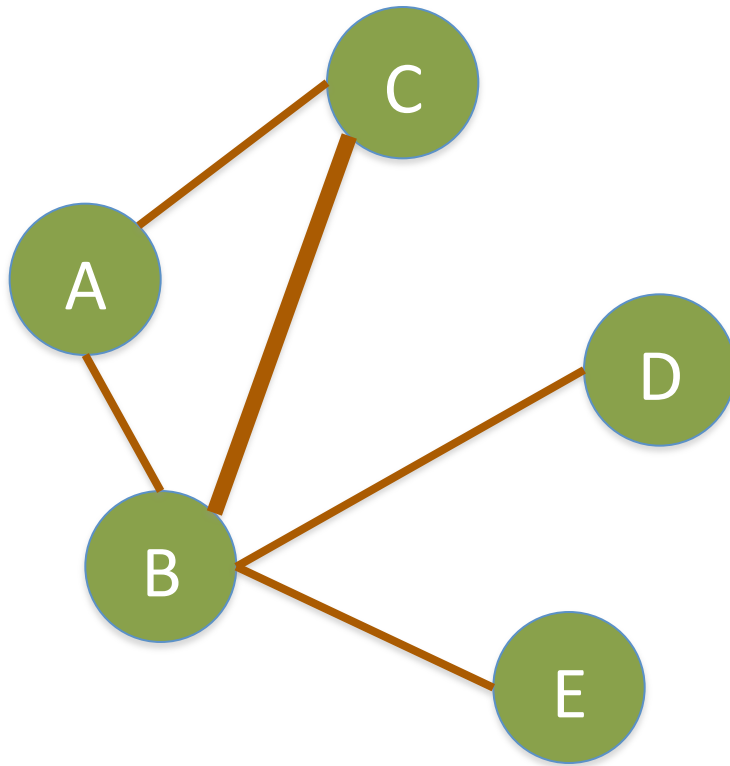
$$B \rightarrow E: 0/1 = 0$$

$$D \rightarrow E: 0/1 = 0$$

$$\text{Total: } \quad \quad \quad \underline{\quad 0 \quad}$$

C: Betweenness Centrality = 0

Betweenness Centrality



A: 0

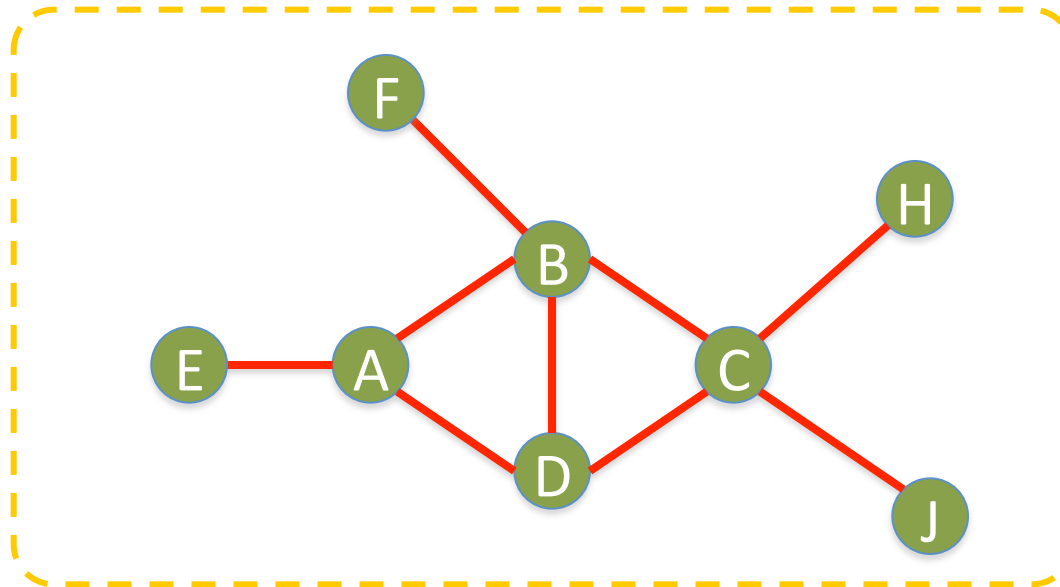
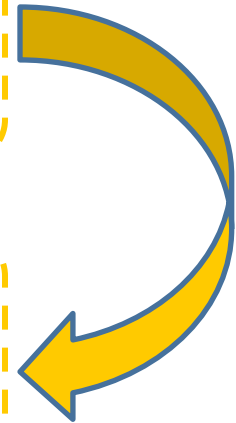
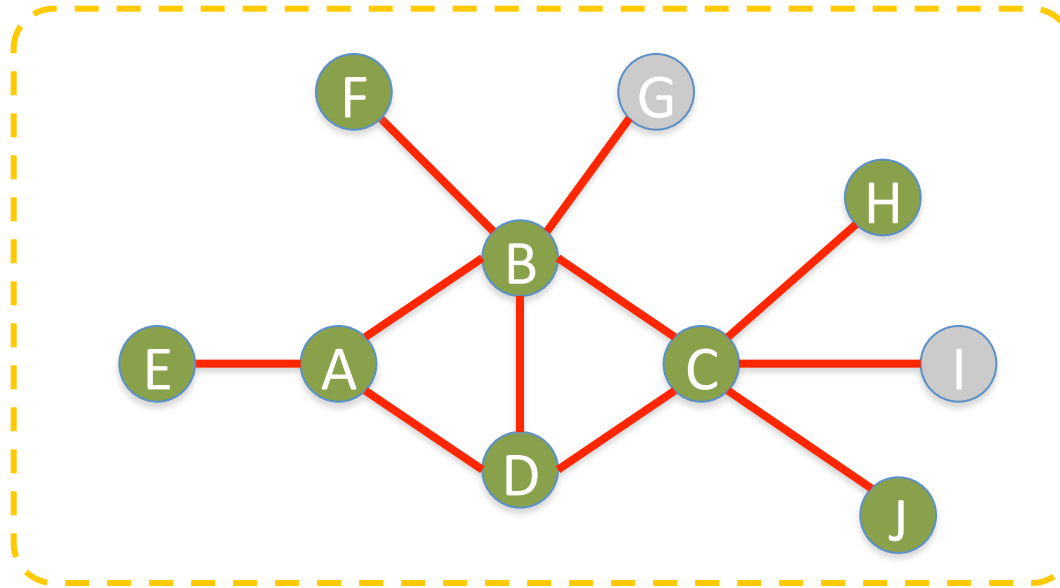
B: 5

C: 0

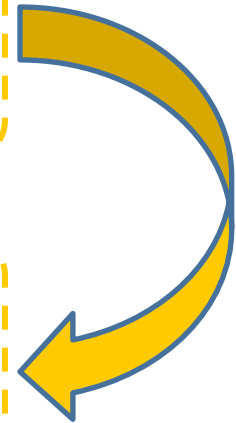
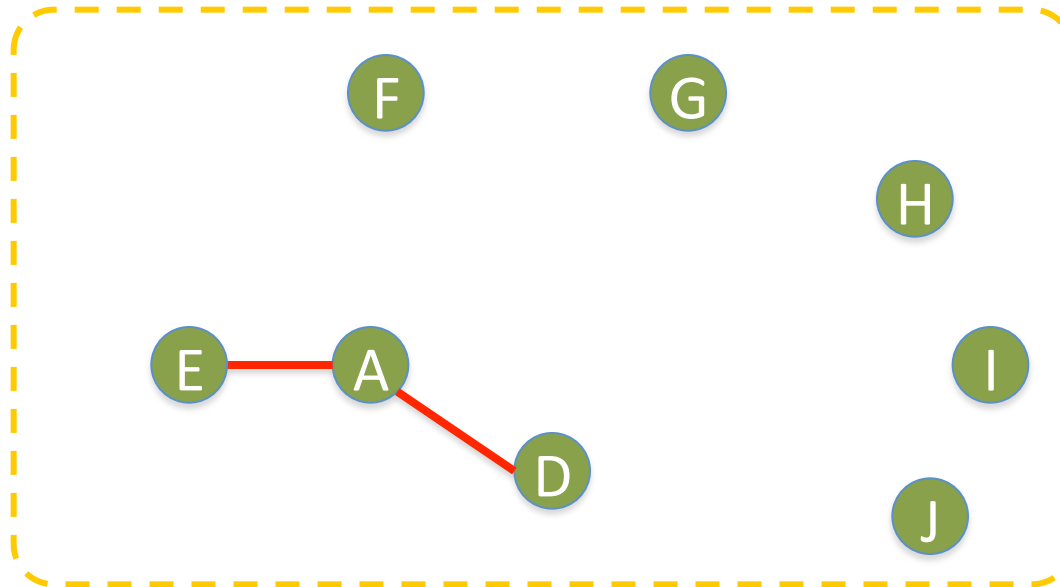
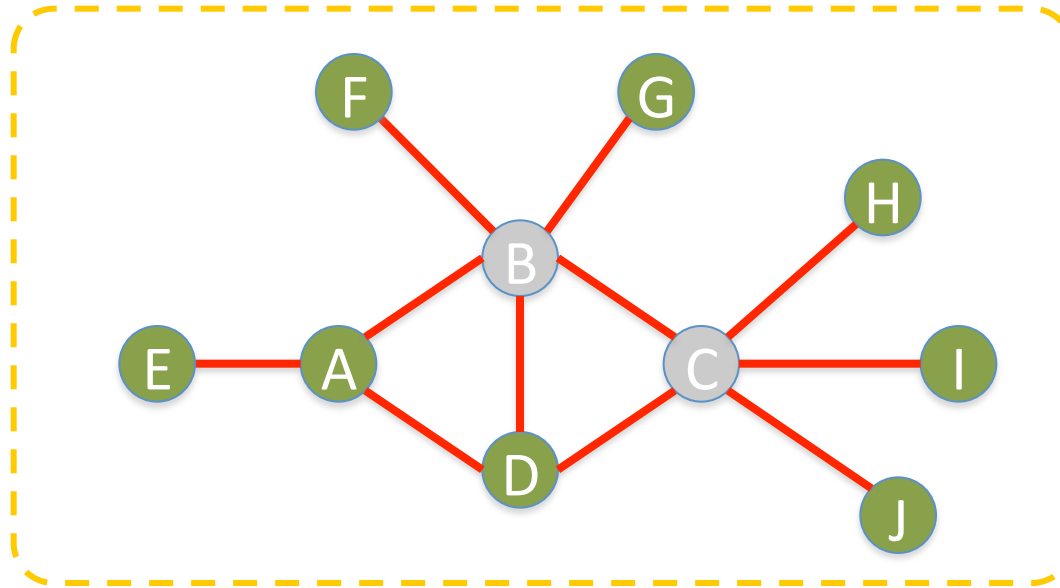
D: 0

E: 0

Which Node is Most **Important**?

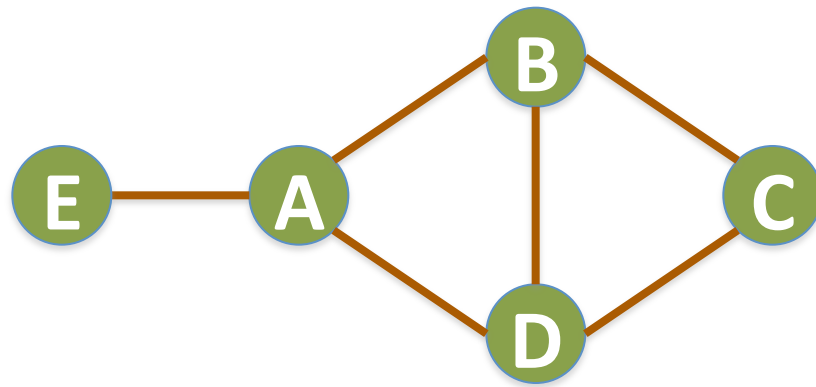


Which Node is Most **Important**?

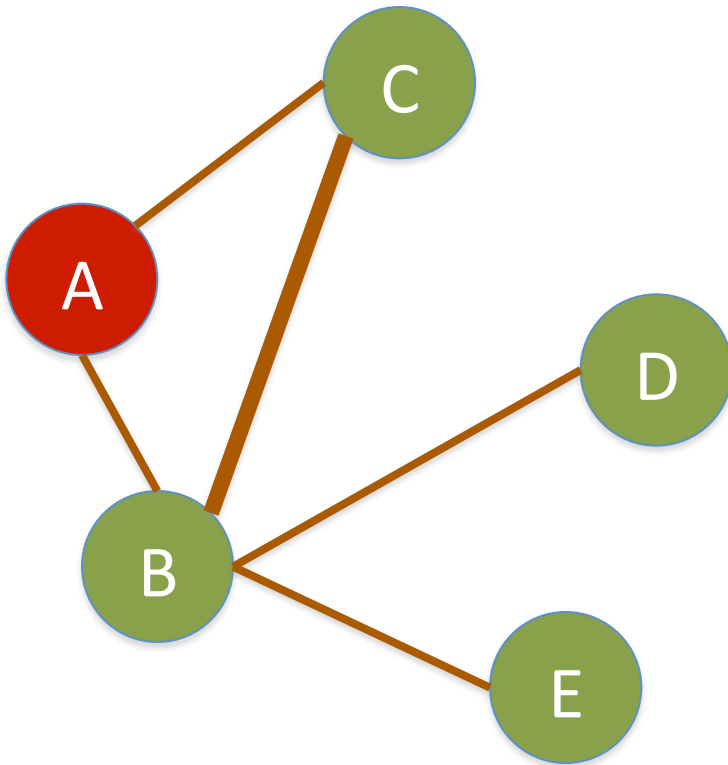


Betweenness Centrality

$$C_B(i) = \sum_{j < k} g_{ik}(i) / g_{jk}$$



Betweenness Centrality



A:

$$B \rightarrow C: 0/1 = 0$$

$$B \rightarrow D: 0/1 = 0$$

$$B \rightarrow E: 0/1 = 0$$

$$C \rightarrow D: 0/1 = 0$$

$$C \rightarrow E: 0/1 = 0$$

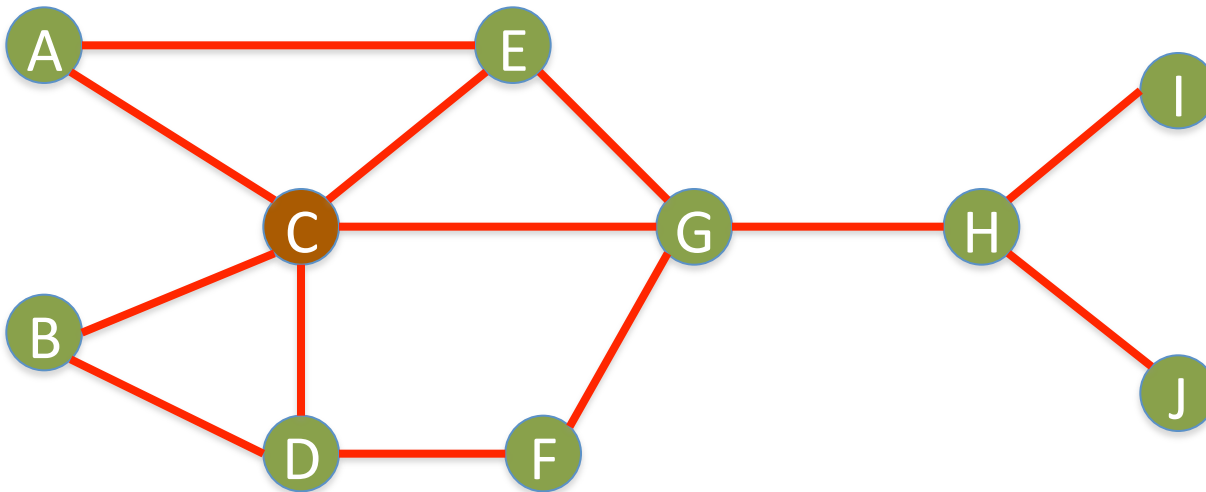
$$D \rightarrow E: 0/1 = 0$$

Total: 0

A: Betweenness Centrality = 0

Closeness
Centrality

Social Network Analysis: Closeness Centrality

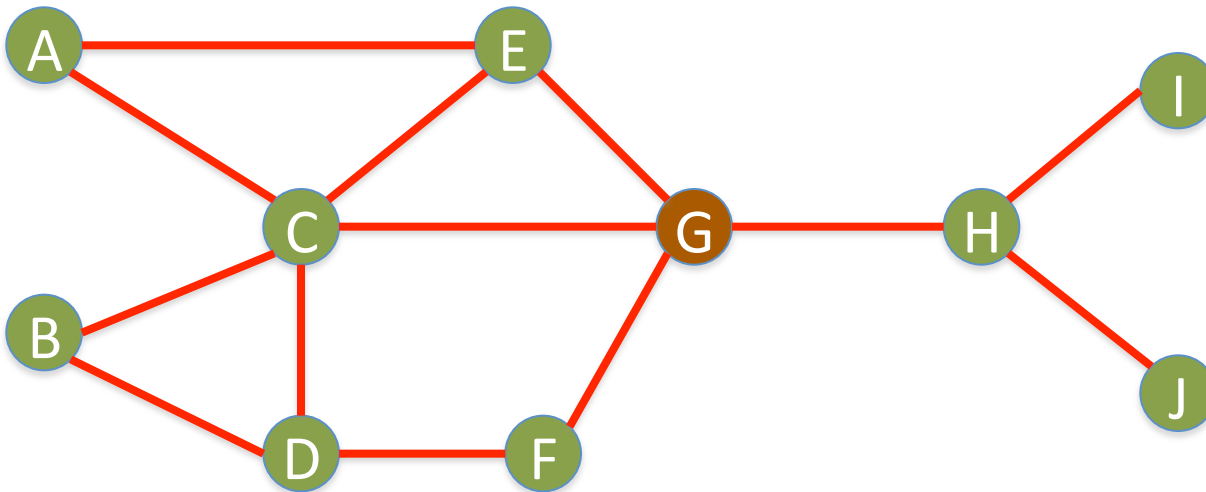


C→A: 1
C→B: 1
C→D: 1
C→E: 1
C→F: 2
C→G: 1
C→H: 2
C→I: 3
C→J: 3

Total=15

C: Closeness Centrality = $15/9 = 1.67$

Social Network Analysis: Closeness Centrality

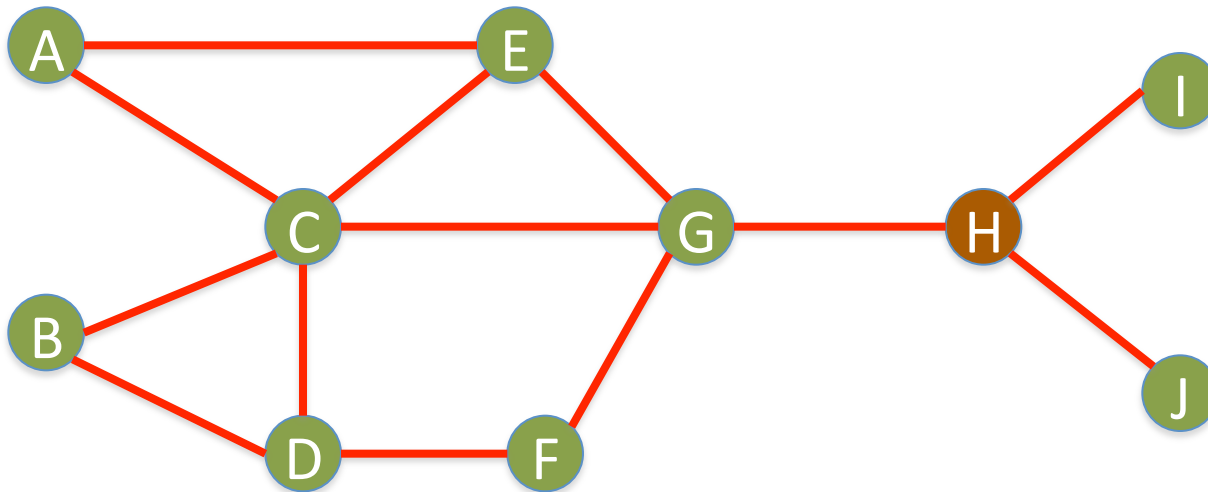


G→A: 2
G→B: 2
G→C: 1
G→D: 2
G→E: 1
G→F: 1
G→H: 1
G→I: 2
G→J: 2

Total=14

G: Closeness Centrality = $14/9 = 1.56$

Social Network Analysis: Closeness Centrality

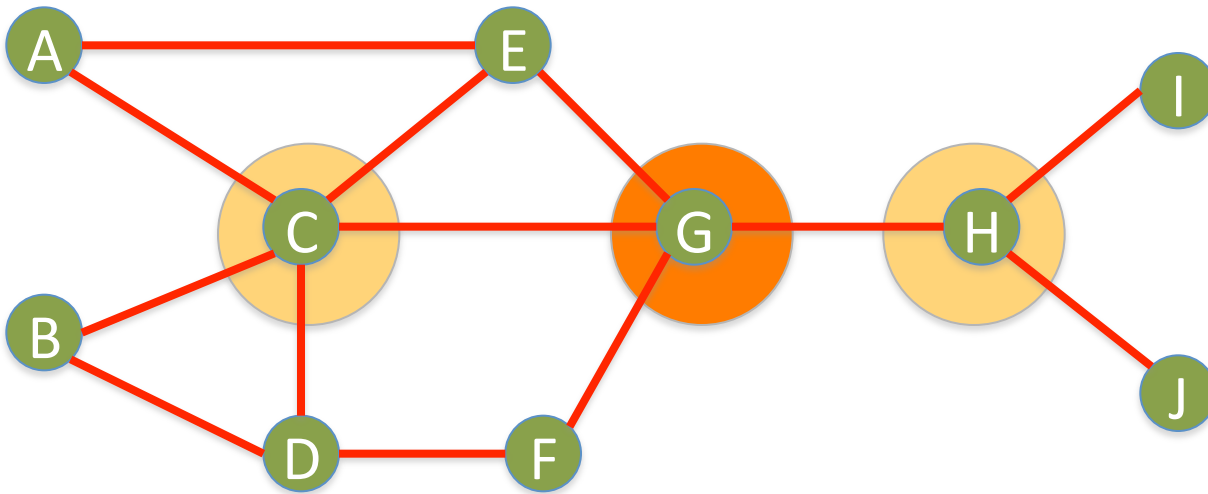


H→A: 3
H→B: 3
H→C: 2
H→D: 2
H→E: 2
H→F: 2
H→G: 1
H→I: 1
H→J: 1

Total=17

H: Closeness Centrality = $17/9 = 1.89$

Social Network Analysis: Closeness Centrality



G: Closeness Centrality = $14/9 = 1.56$ ①

C: Closeness Centrality = $15/9 = 1.67$ ②

H: Closeness Centrality = $17/9 = 1.89$ ③

Eigenvector centrality:

Importance of a node
depends on
the importance of its neighbors

Social Network Analysis: Closeness Centrality

Sum of the reciprocal distances

$$C_C(p_k) = \sum_{i=1}^n d(p_i, p_k)^{-1}$$

where $d(p_j, p_k)$ is the geodesic distance (shortest paths) linking p_j, p_k

Social Network Analysis: Betweenness Centrality

$$C_B(p_k) = \sum_{i < j}^n \frac{g_{ij}(p_k)}{g_{ij}}; \quad i \neq j \neq k$$

where g_{ij} is the geodesic distance (shortest paths) linking p_i and p_j and $g_{ij}(p_k)$ is the geodesic distance linking p_i and p_j that contains p_k .

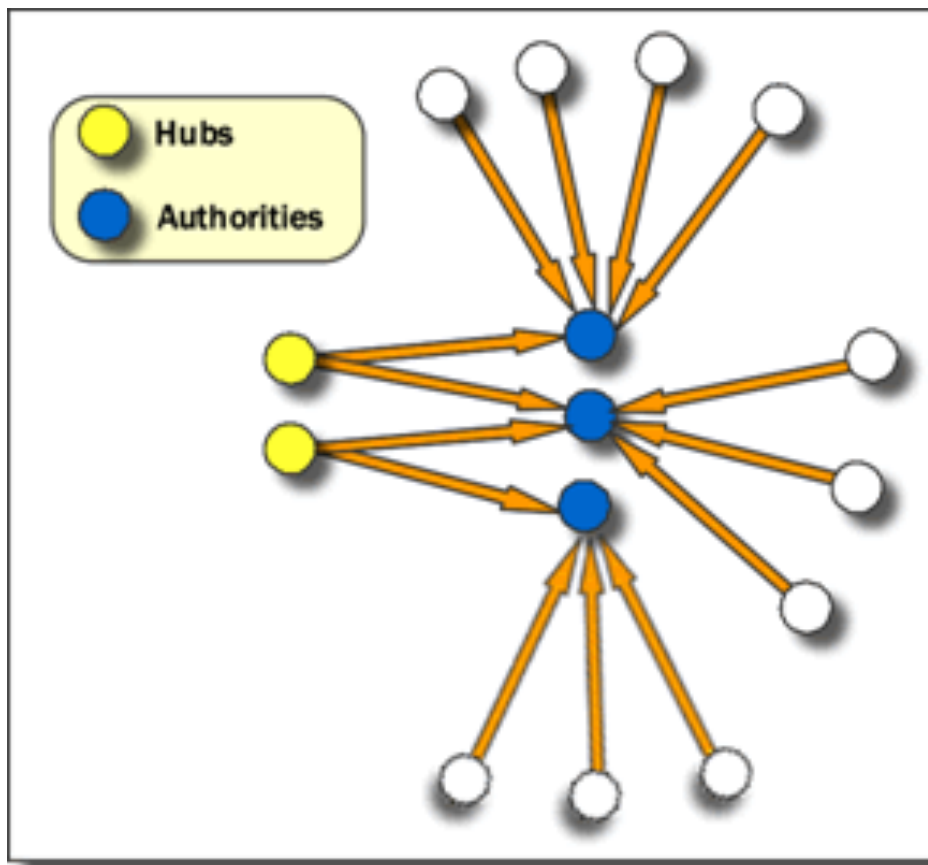
Social Network Analysis: Degree Centrality

$$C_D(p_k) = \sum_{i=1}^n a(p_i, p_k)$$

where $a(p_i, p_k) = 1$ if and only if p_i and p_k are connected by a line
0 otherwise

$$C'_D(p_k) = \frac{\sum_{i=1}^n a(p_i, p_k)}{n-1}$$

Social Network Analysis: Hub and Authority



Hubs are entities that point to a relatively large number of authorities. They are essentially the mutually reinforcing analogues to authorities. Authorities point to high hubs. Hubs point to high authorities. You cannot have one without the other.

Application of SNA

**Social Network Analysis
of
Research Collaboration
in
Information Reuse and Integration**

Example of SNA Data Source


















dblp
computer science bibliography













home | browse | search | about



IRI 2010: Las Vegas, NV, USA

-    **Proceedings of the IEEE International Conference on Information Reuse and Integration, IRI 2010, 4-6 August 2010, Las Vegas, Nevada, USA.** IEEE Systems, Man, and Cybernetics Society 2010
-    Reda Alhajj, James B. D. Joshi, Mei-Ling Shyu: **Message from Program Co-Chairs.** 1
-    Stuart Harvey Rubin, Shu-Ching Chen: **Forward.** 1
-    Lotfi A. Zadeh: **Precisiation of meaning - toward computation with natural language.** 1-4
-    Reda Alhajj, Shu-Ching Chen, Gongzhu Hu, James B. D. Joshi, Gordon K. Lee, Stuart Harvey Rubin, Mei-Ling Shyu, Lotfi A. Zadeh: **Panel title: Critical need for funding of basic and applied research in large-scale computing.** 1

Automation, Integration and Reuse across Various Apps

-    László István Etesi, André Csillaghy, Lin-Ching Chang: **A message-based interoperability framework with application to astrophysics.** 1-6
-    Awny Alnusair, Tian Zhao, Eric Bodden: **Effective API navigation and reuse.** 7-12
-    Manabu Ohta, Ryohei Inoue, Atsuhiko Takasu: **Empirical evaluation of active sampling for CRF-based analysis of pages.** 13-18
-    Qunzhi Zhou, Viktor K. Prasanna: **Workflow management of simulation based computation processes in transportation domain.** 19-24

Source: <http://www.informatik.uni-trier.de/~ley/db/conf/iri/iri2010.html>

Research Question

- RQ1: What are the scientific **collaboration patterns** in the IRI research community?
- RQ2: Who are the **prominent researchers** in the IRI community?

Methodology

- Developed a simple **web focused crawler** program to download literature information about all IRI papers published between **2003 and 2010** from **IEEE Xplore** and **DBLP**.
 - **767** paper
 - **1599** distinct author
- Developed a program to convert the list of coauthors into the **format of a network file** which can be readable by social network analysis software.
- **UCINet** and **Pajek** were used in this study for the social network analysis.

Top10 prolific authors (IRI 2003-2010)

1. Stuart Harvey Rubin
2. Taghi M. Khoshgoftaar
3. Shu-Ching Chen
4. Mei-Ling Shyu
5. Mohamed E. Fayad
6. Reda Alhajj
7. Du Zhang
8. Wen-Lian Hsu
9. Jason Van Hulse
10. Min-Yuh Day

Data Analysis and Discussion

- **Closeness Centrality**
 - Collaborated widely
- **Betweenness Centrality**
 - Collaborated diversely
- **Degree Centrality**
 - Collaborated frequently
- **Visualization of Social Network Analysis**
 - Insight into the structural characteristics of research collaboration networks

Top 20 authors with the highest **closeness** scores

Rank	ID	Closeness	Author
1	3	0.024675	Shu-Ching Chen
2	1	0.022830	Stuart Harvey Rubin
3	4	0.022207	Mei-Ling Shyu
4	6	0.020013	Reda Alhajj
5	61	0.019700	Na Zhao
6	260	0.018936	Min Chen
7	151	0.018230	Gordon K. Lee
8	19	0.017962	Chengcui Zhang
9	1043	0.017962	Isai Michel Lombera
10	1027	0.017962	Michael Armella
11	443	0.017448	James B. Law
12	157	0.017082	Keqi Zhang
13	253	0.016731	Shahid Hamid
14	1038	0.016618	Walter Z. Tang
15	959	0.016285	Chengjun Zhan
16	957	0.016285	Lin Luo
17	956	0.016285	Guo Chen
18	955	0.016285	Xin Huang
19	943	0.016285	Sneh Gulati
20	960	0.016071	Sheng-Tun Li

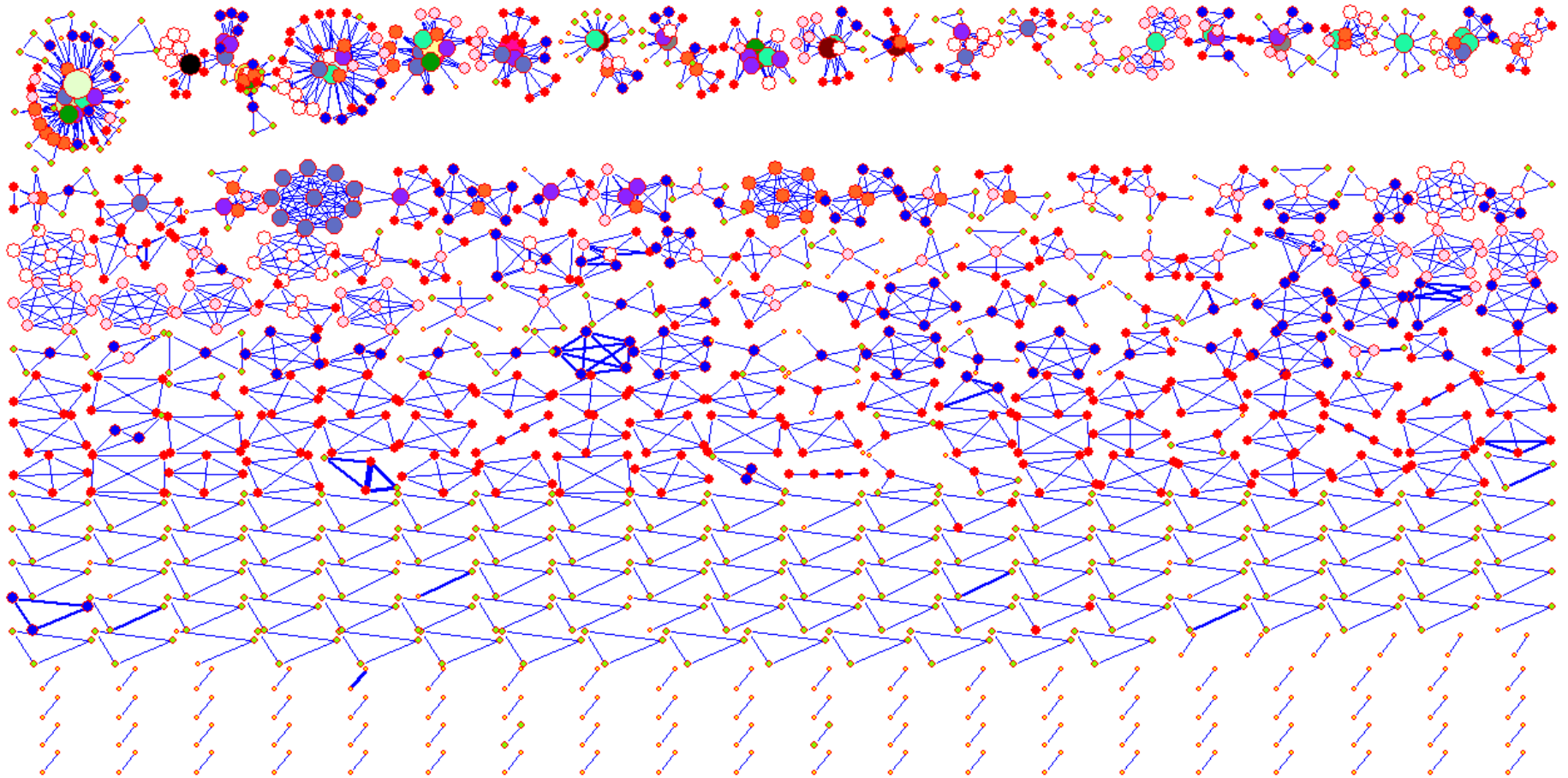
Top 20 authors with the highest **betweenness** scores

Rank	ID	Betweenness	Author
1	1	0.000752	Stuart Harvey Rubin
2	3	0.000741	Shu-Ching Chen
3	2	0.000406	Taghi M. Khoshgoftaar
4	66	0.000385	Xingquan Zhu
5	4	0.000376	Mei-Ling Shyu
6	6	0.000296	Reda Alhajj
7	65	0.000256	Xindong Wu
8	19	0.000194	Chengcui Zhang
9	39	0.000185	Wei Dai
10	15	0.000107	Narayan C. Debnath
11	31	0.000094	Qianhui Althea Liang
12	151	0.000094	Gordon K. Lee
13	7	0.000085	Du Zhang
14	30	0.000072	Baowen Xu
15	41	0.000067	Hongji Yang
16	270	0.000060	Zhiwei Xu
17	5	0.000043	Mohamed E. Fayad
18	110	0.000042	Abhijit S. Pandya
19	106	0.000042	Sam Hsu
20	8	0.000042	Wen-Lian Hsu

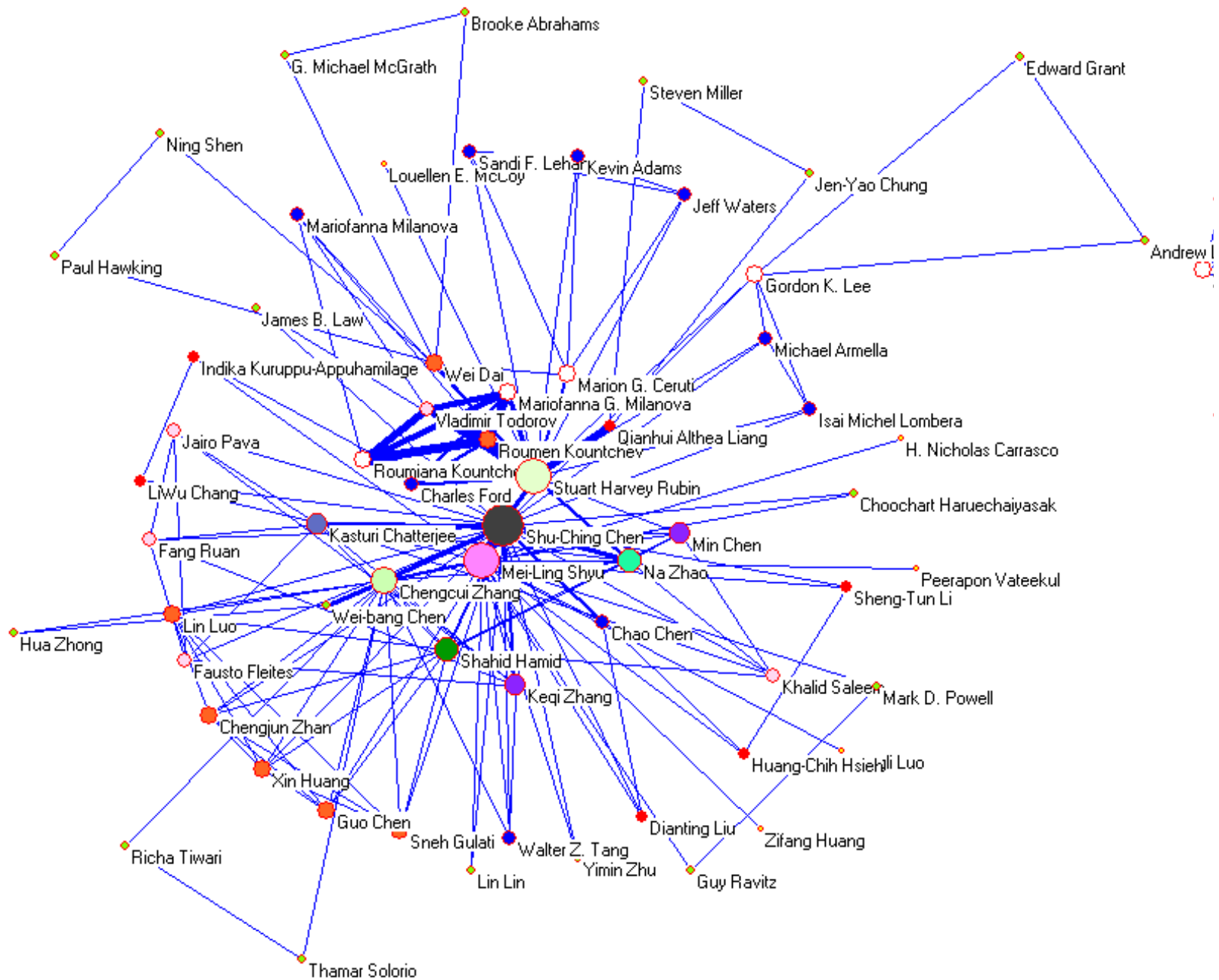
Top 20 authors with the highest **degree** scores

Rank	ID	Degree	Author
1	3	0.035044	Shu-Ching Chen
2	1	0.034418	Stuart Harvey Rubin
3	2	0.030663	Taghi M. Khoshgoftaar
4	6	0.028786	Reda Alhajj
5	8	0.028786	Wen-Lian Hsu
6	10	0.024406	Min-Yuh Day
7	4	0.022528	Mei-Ling Shyu
8	17	0.021277	Richard Tzong-Han Tsai
9	14	0.017522	Eduardo Santana de Almeida
10	16	0.017522	Roumen Kountchev
11	40	0.016896	Hong-Jie Dai
12	15	0.015645	Narayan C. Debnath
13	9	0.015019	Jason Van Hulse
14	25	0.013767	Roumiana Kountcheva
15	28	0.013141	Silvio Romero de Lemos Meira
16	24	0.013141	Vladimir Todorov
17	23	0.013141	Mariofanna G. Milanova
18	5	0.013141	Mohamed E. Fayad
19	19	0.012516	Chengcui Zhang
20	18	0.011890	Waleed W. Smari

Visualization of IRI (IEEE IRI 2003-2010) co-authorship network (global view)

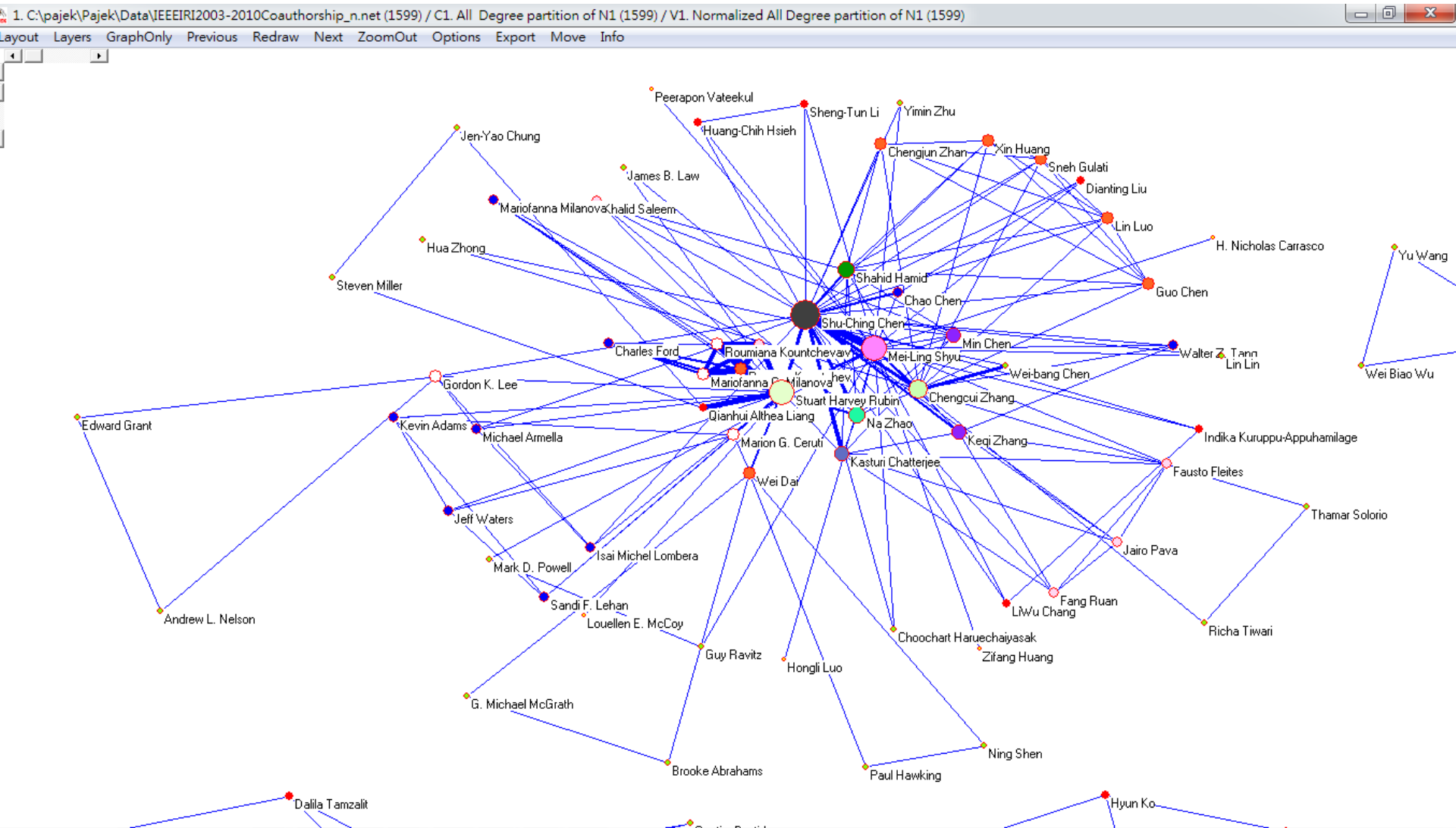


Source: Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011),
"Social Network Analysis of Research Collaboration in Information Reuse and Integration"



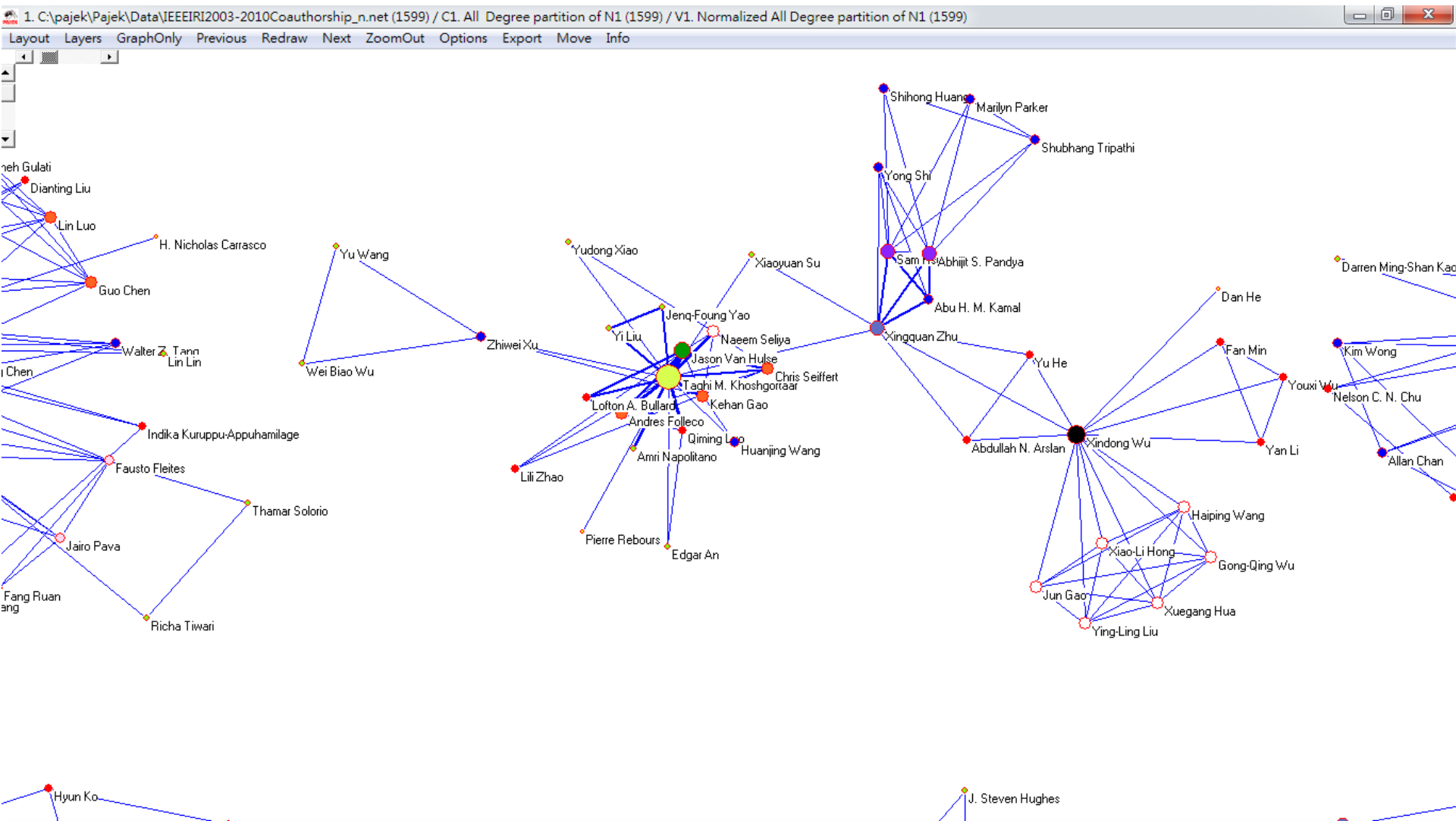
Source: Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011),
 "Social Network Analysis of Research Collaboration in Information Reuse and Integration"

Visualization of Social Network Analysis



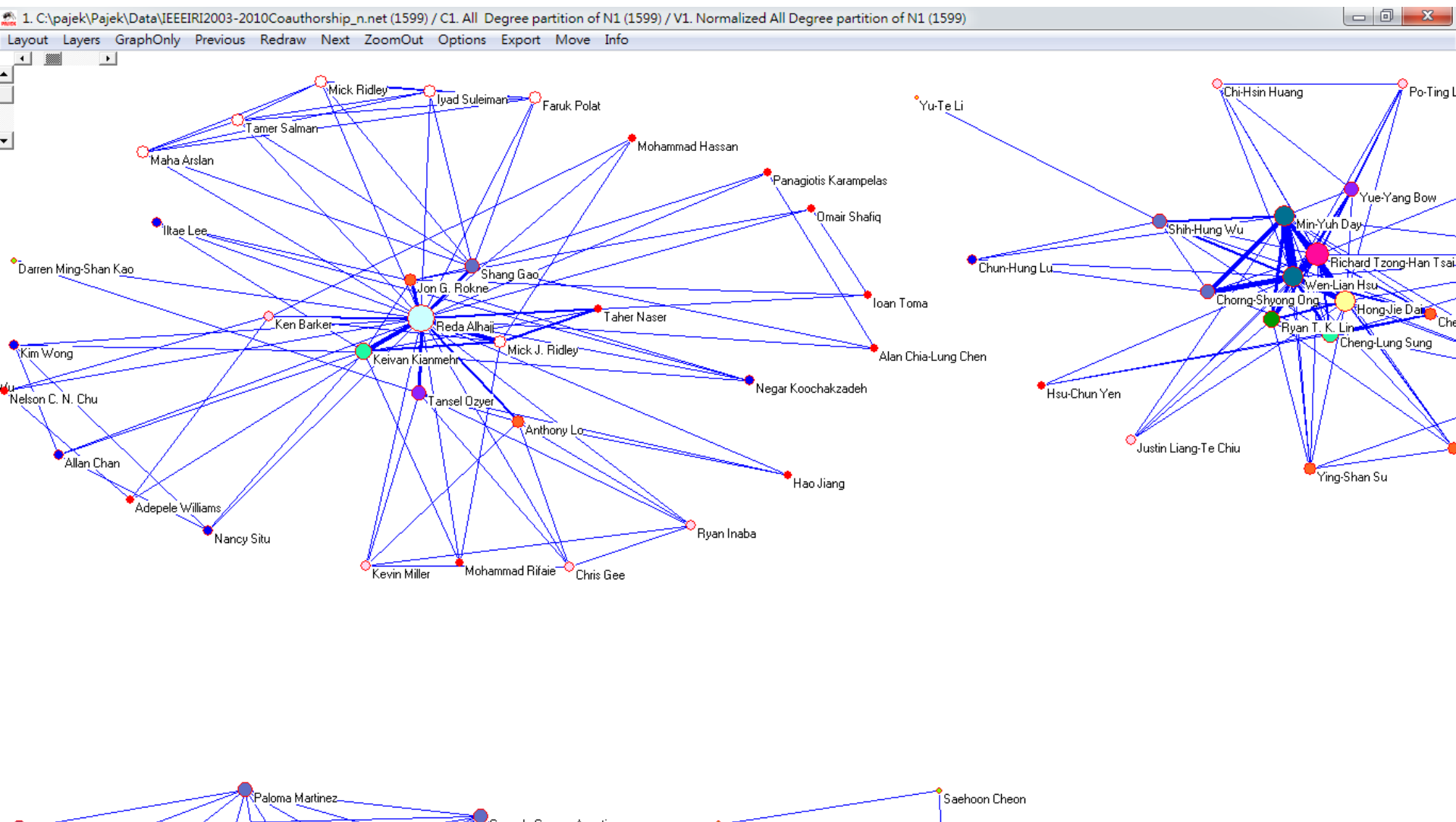
Source: Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011),
"Social Network Analysis of Research Collaboration in Information Reuse and Integration"

Visualization of Social Network Analysis



Source: Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011),
"Social Network Analysis of Research Collaboration in Information Reuse and Integration"

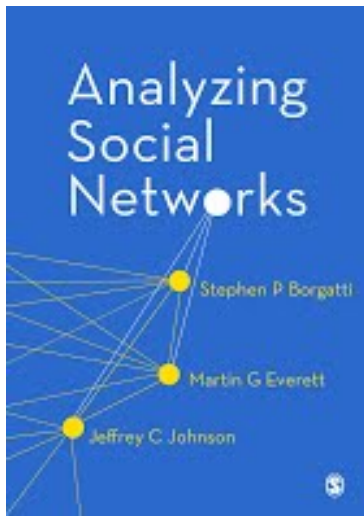
Visualization of Social Network Analysis



Source: Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011),
"Social Network Analysis of Research Collaboration in Information Reuse and Integration"

Social Network Analysis (SNA) Tools

- **NetworkX**
- **igraph**
- **Gephi**
- **UCINet**
- **Pajek**



NetworkX

NetworkX

[NetworkX Home](#) | [Documentation](#) | [Download](#) | [Developer \(Github\)](#)

High-productivity software for complex networks

NetworkX is a Python language software package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.



[Documentation](#)

all documentation

[Examples](#)

using the library

[Reference](#)

all functions and methods

Features

- Python language data structures for graphs, digraphs, and multigraphs.
- Many standard graph algorithms
- Network structure and analysis measures
- Generators for classic graphs, random graphs, and synthetic networks
- Nodes can be "anything" (e.g. text, images, XML records)
- Edges can hold arbitrary data (e.g. weights, time-series)
- Open source [BSD license](#)
- Well tested: more than 1800 unit tests, >90% code coverage
- Additional benefits from Python: fast prototyping, easy to teach, multi-platform

Versions

Latest Release

networkx-1.11

30 January 2016

[downloads](#) | [docs](#) | [pdf](#)

Development

2.0dev

[github](#) | [docs](#) | [pdf](#)

build passing

coverage 94%

Contact

[Mailing list](#)
[Issue tracker](#)



<https://networkx.github.io/>

igraph



Products ▾

News

On github



igraph – The network analysis package

igraph is a collection of network analysis tools with the emphasis on **efficiency, portability** and ease of use. igraph is **open source** and free. igraph can be programmed in **R, Python** and **C/C++**.

igraph R package

python-igraph

igraph C library

R/igraph 1.0.0

Repositories at Github

R/igraph 0.7.1

C/igraph 0.7.1

R/igraph 0.7.0

python-igraph 0.7.0

C/igraph 0.7.0

Recent news

R/igraph 1.0.0

June 24, 2015

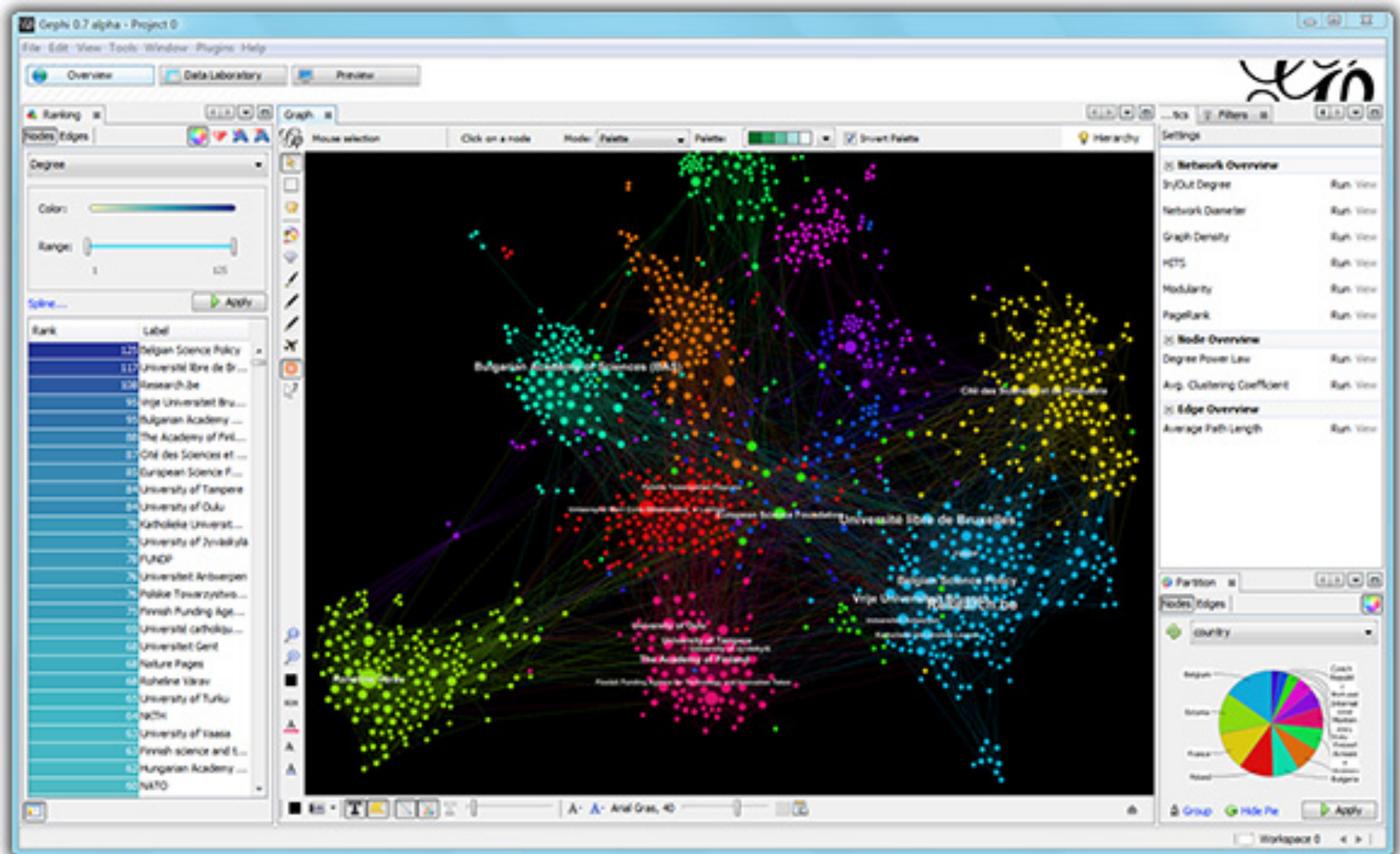
Release Notes

This is a new major release, with a lot of UI changes. We tried to make it easier to

<http://igraph.org/>

Gephi

The Open Graph Viz Platform



SNA Tool: UCINET

<https://sites.google.com/site/ucinetsoftware/home>



UCINET Software

Search this site

[Home](#) [Download](#) [Versions](#) [FAQs & Tips](#) [Buy!](#)

Navigation

- [Home](#)
- [Datasets](#)
- [Download](#)
- [History](#)
- ▼ [How to Use](#)
 - [QuickStart guide](#)
 - [Users group](#)
 - [FAQs & Tips](#)
 - [Tutorials](#)
 - [Spanish Resources](#)
 - [Workshops](#)
 - [More ...](#)
- ▼ [Purchasing](#)
 - [Info](#)
 - [Buy now](#)
 - [Trial Version](#)
- ▼ [Versions](#)
 - [Recent](#)
 - [Jan 2012 only](#)
 - [2011 and before](#)
 - [Fix List](#)

Links

- ▼ [Analytic Technologies](#)
 - [NetDraw](#)
 - [E-Net](#)
- ▼ [LINKS Center](#)

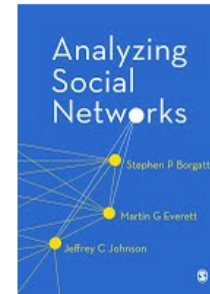
New! UCINET-oriented book on social network analysis now available! See [details](#).

UCINET 6 for Windows is a software package for the analysis of social network data. It was developed by Lin Freeman, Martin Everett and Steve Borgatti. It comes with the NetDraw network visualization tool.

If you use the software, please cite it. Here is a sample citation:

- **Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.**

For customer support (e.g., ordering info, billing etc) contact roberta@analytictech.com. For tech support join the [users group](#) or contact support@analytictech.com. We prefer you try the users group first since the answer to your question may benefit others.



Download and/or Purchase

- The program can be [downloaded](#) and used for free for 90 days. In addition, students can [purchase](#) the downloaded program for \$40. Faculty and government can purchase the downloaded program for \$150, and all others pay \$250. Site licenses and extremely generous volume discounts are available.
- Note that all purchases are provided as electronic downloads. If necessary you can order a CD from us for an exorbitant fee, but there is no reason to do this. Purchasers of the software are welcome to burn their own CDs at will. They are also free to download the program to all of their computers.
- For more details, including questions about taxes, shipping costs, payment methods, etc., please visit the [Order Info](#) page.

News

[Problem with NetDraw](#) The current UCINET installation package includes NetDraw version 2.141. Unfortunately, this version of NetDraw has a problem reading UCINET system files, and we are temporarily unable to update the ...
Posted Dec 12, 2014, 1:22 PM by Steve Borgatti

Showing posts 1 - 1 of 7.
[View more »](#)

Current Version

[Version 6.538 | 12 Dec 2014](#) This version replaces the old "spreadsheet editor" with a new "matrix editor", which is available from

SNA Tool: Pajek

<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>



Networks / Pajek

Program for Large Network Analysis

In January 2008 this page was replaced by [Pajek Wiki](#).

Pajek runs on Windows and is free for noncommercial use.

[DOWNLOAD Pajek](#)

Data: [test networks](#), [GPHs](#), [GEDs](#), [PDB files](#).

[Screenshots](#); [History](#); [Manual \(pdf\)](#); [Papers/presentations](#); [Applications](#); [in News](#); [Examples](#): [SVG](#), [PDF](#).

[How to ?](#) [English](#) / [Slovene](#) / [Japanese](#) (problems with IE - download and use Acrobat reader).

[Pajek nicely runs on Linux via Wine](#), [Converting Excel/text into Pajek format](#).

[Pajek to SVG animation](#), [WoS to Pajek](#).

[Slides from NICTA workshop](#), Sydney, Australia, June 14-17, 2005.

[Slides from workshop at GD'05](#), Limerick, Ireland, Sept 11-14, 2005.

[Pajek workshop at XXVIII Sunbelt Conference](#), St. Pete Beach, Florida, USA, January 22-27, 2008: [slides](#).

[Network analysis course at ECPR Summer School in Methods and Techniques](#), Ljubljana, Slovenia, July 30 - August 16, 2008.

W. de Nooy, A. Mrvar, V. Batagelj: *Exploratory Social Network Analysis with Pajek*, CUP, January 2005; [ESNA page](#).

P. Doreian, V. Batagelj, A. Ferligoj: *Generalized Blockmodeling*, CUP, November 2004.

Chapter about Pajek: V. Batagelj, A. Mrvar: *Pajek - Analysis and Visualization of Large Networks*.

in Jünger, M., Mutzel, P., (Eds.) *Graph Drawing Software*. Springer, Berlin 2003. p. 77-103 / [Amazon](#).

An improved version of the paper presented at [Sunbelt'97](#) was published in [Connections](#) 21(1998)2, 47-57 - V. Batagelj, A. Mrvar: *Pajek - Program for Large Network Analysis* (PDF; [PRISON.KIN](#)).

Our layouts for *Graph-Drawing Competitions*: [GD95](#), [GD96](#), [GD97](#), [GD98](#), [GD99](#), [GD00](#), [GD01](#) and [GD05](#).

[Mladina](#) (front page): [Paiek in Koeln](#); [PaiekMan](#) in [Osoie \(Ossiach, Austria\)](#):

SNA Tool: Pajek

<http://pajek.imfm.si/doku.php>

PAJEK WIKI



navigation

- [Pajek](#)
- [News](#)
- [Network Analysis](#)
- [How to](#)
- [Kako](#)
- [Download](#)
- [Network data](#)
- [Events](#)
- [Resources](#)
- [Links](#)
- [User's Comments](#)
- [Pajek mailing list](#)
- [IMFM ePrints](#)

search

toolbox

- [Sitemap](#)
- [Recent changes](#)
- [Backlinks](#)
- [Login](#)

[Show pagesource](#)

[Old revisions](#)

Trace: • start

Pajek Wiki

This Wiki is about program [Pajek](#) and (large) network analysis and visualization.

Most of its content is provided by authors of Pajek and the Wiki users can only read it.

The logged-in users can create/edit pages in [Questions and Answers](#) area and in [User's Comments](#) area. They can also join the [Pajek mailing list](#) and discuss their ideas or problems on it.

To login you need a **user name** and a **password**. To get them you click on login in the toolbox and then register - the password will be sent to you by e-mail.

The rules for writing Wiki pages are described on [Syntax](#) page.

Some trusted users will be promoted to **writers**. They will be allowed to create/edit on pages with

edit: **writer**

at the bottom of the page - such as for example [events](#), [introduction](#) and [selected topics](#).

start.txt · Last modified: 2008/05/23 12:42 by vlado

[Back to top](#)

CHIMERIC DE

LICENSED

W3C CSS


DOKUWIKI

GET FIREFOX

RSS XML FEED

W3C XHTML 1.0

About Pajek



Pajek 0.73
 November 1996 - August 2001
 Copyright (c) 1996
 Vladimir Batagelj and Andrej Mrvar
 All rights reserved.

Free for noncommercial use.
<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

Report

File

2. C:\PAJEK\DATA\RAZNO\imrich.net [2-Mode] (674)

Number of vertices (n): 674
 Number of arcs: 0
 0 loops
 Number of edges: 613
 0 loops
 Density1 [loops allowed] = 0.0026988
 Density2 [no loops allowed] = 0.0027028

2. Affiliation partition of N2 [314,360] (674)

Dimension: 674



Recycle Bin



Pajek

Pajek

File Net Nets Operations Partition Partitions Permut. Cluster Hierarchy Vector Vectors Options Draw Macro Info

Net: Transform (selected)

- Random Network
- Partitions
- Components
- Hierarchical Decomposition
- Numbering
- Citation Weights
- k-Neighbours
- Paths between 2 vertices
- Critical Path Method - CPM
- Maximum Flow
- Vector

Partition:

- Transpose
- Remove
- Add
- Edges->Arcs
- Arcs->Edges
- Reduction
- Generate in Time
- 2-Mode to 1-Mode (selected)
 - Rows
 - Columns
 - Include Loops
 - Multiple Lines
 - Normalize 1-Mode (selected)
 - Geo (selected)
 - Input
 - Output
 - Min
 - Max
 - MinDir
 - MaxDir
- Sort Lines

Partitions:

- t [2-Mode] (674)
- 674)

Permut.:

- 2-Mode to 1-Mode

Cluster:

- 2-Mode to 1-Mode

Hierarchy: [Empty field]

Vector: [Empty field]

References

- Bing Liu (2011) , “Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data,” 2nd Edition, Springer.
<http://www.cs.uic.edu/~liub/WebMiningBook.html>
- Jennifer Golbeck (2013), Analyzing the Social Web, Morgan Kaufmann.
<http://analyzingthesocialweb.com/course-materials.shtml>
- Sentinel Visualizer, <http://www.fmsasg.com/SocialNetworkAnalysis/>
- Min-Yuh Day, Sheng-Pao Shih, Weide Chang (2011), "Social Network Analysis of Research Collaboration in Information Reuse and Integration," The First International Workshop on Issues and Challenges in Social Computing (WICSOC 2011), August 2, 2011, in Proceedings of the IEEE International Conference on Information Reuse and Integration (IEEE IRI 2011), Las Vegas, Nevada, USA, August 3-5, 2011, pp. 551-556.