The New KONECT Project

at the University of Namur

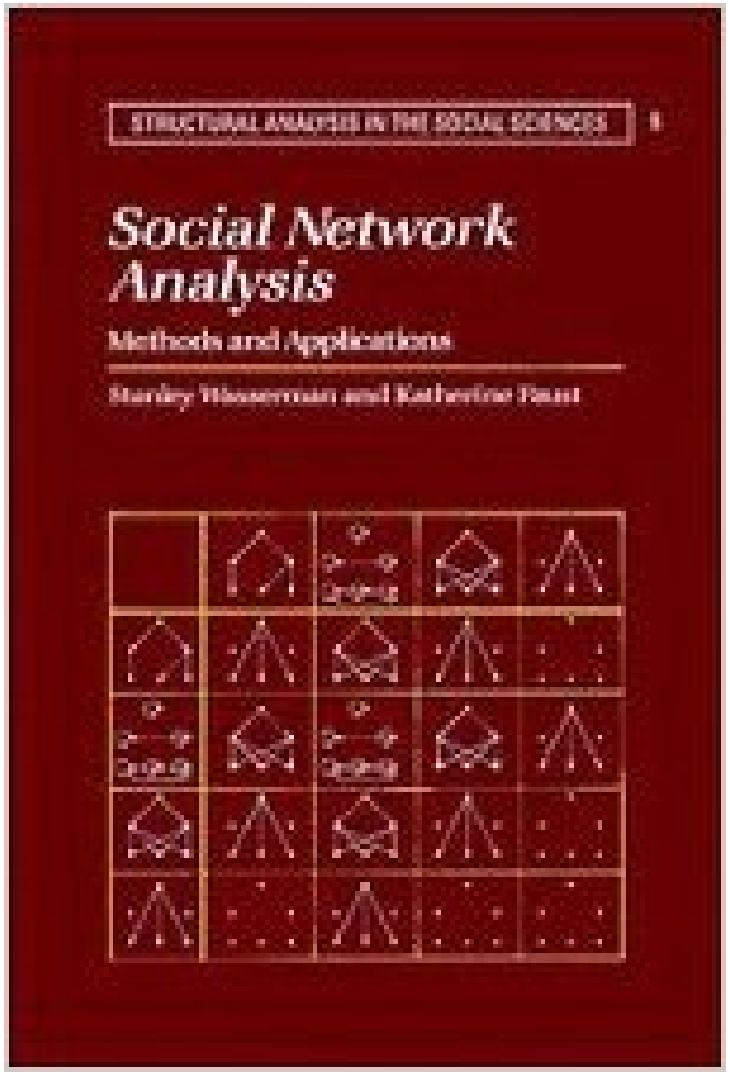
Jérôme KUNEGIS – with acknowledgments to many people

and a short introduction to Stu

Table B.3. *"Reports to" relation between managers of Krackhardt's high-tech company*

Manager	
1	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
4	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
5	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
6	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
7	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
8	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
9	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
10	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
11	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
12	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
13	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
14	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
15	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
16	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
17	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
18	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
19	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
20	0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
21	0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Social Network Analysis without the Web

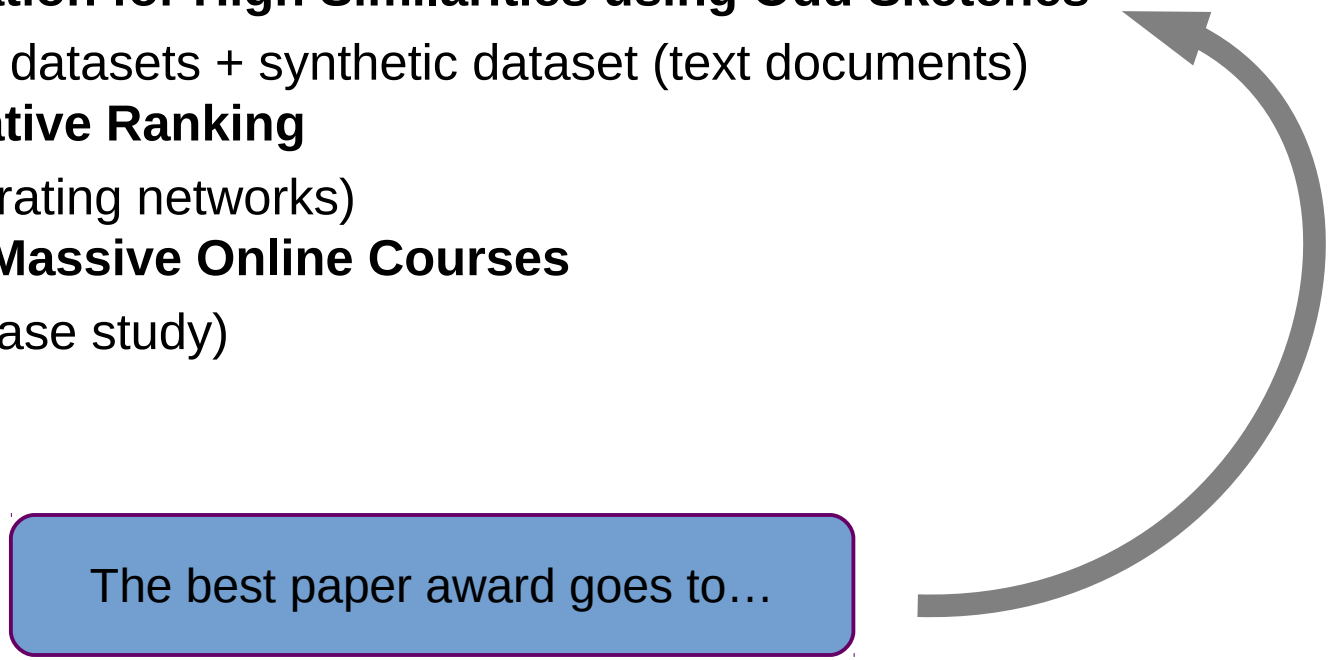


“Social Network Analysis” by
Wasserman & Faust first edition 1994

Contains 18 datasets (based on 19th printing)

WWW 2014 Best Paper Nominations

- **Community-Based Bayesian Aggregation Models for Crowdsourcing**
  - **4** datasets (crowdsourcing)
- **Efficient Estimation for High Similarities using Odd Sketches**
  - **5** real-world datasets + synthetic dataset (text documents)
- **Local Collaborative Ranking**
  - **3** datasets (rating networks)
- **Engaging with Massive Online Courses**
  - **1** dataset (case study)

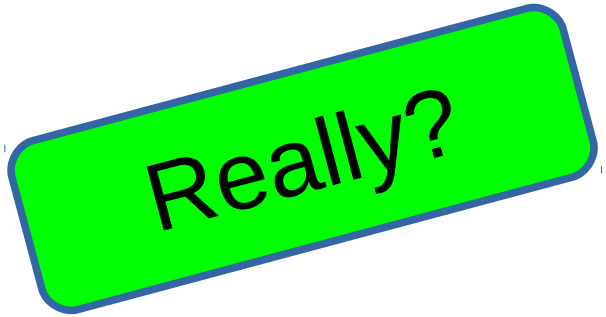


Why Do Researchers Use Multiple Datasets?

- To cover more application areas
- To show that results are generalizable
- To make results more statistically significant

Showing that Algorithm X is Better Than Y

- Setup: We want to compare two prediction algorithms X and Y
- Experiment: Apply X and Y to dataset A
- Result: X has higher precision than Y
- Conclusion: “Algorithm X performs better than algorithm Y”



Let's Make More Experiments

- Add a dataset B to the experiments
- On dataset B, Y performs better than X

More Datasets

- Add more datasets to the mix
- Algorithm X is better than algorithm Y with 6 out of 10 datasets
- Under Null hypothesis of equal probability of X performing better than Y on any one dataset and results on datasets being independent, the probability that this happens is 17%, i.e., not significant!
- Need about 65 datasets to get a statistically significant result at ($p \leq 0.05$) for a 60% result.

The KONECT Project – Koblenz Network Collection


© P!ELmedia

On the Spectral Evolution of Large Networks

Jérôme Kunegis

Institute for Web Science and Technologies
University of Koblenz-Landau
kunegis@uni-koblenz.de

November 2011

Vom Promotionsausschuss des Fachbereichs 4: Informatik der Universität Koblenz-Landau zur Verleihung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation.

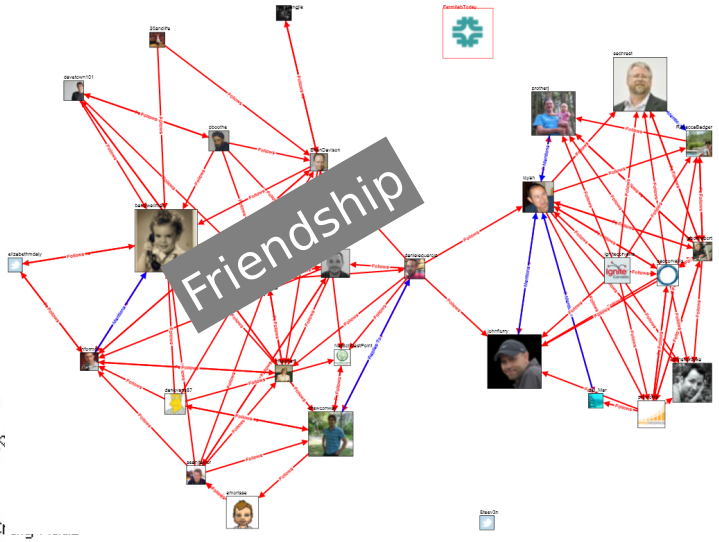
PhD thesis at the University of Koblenz-Landau.

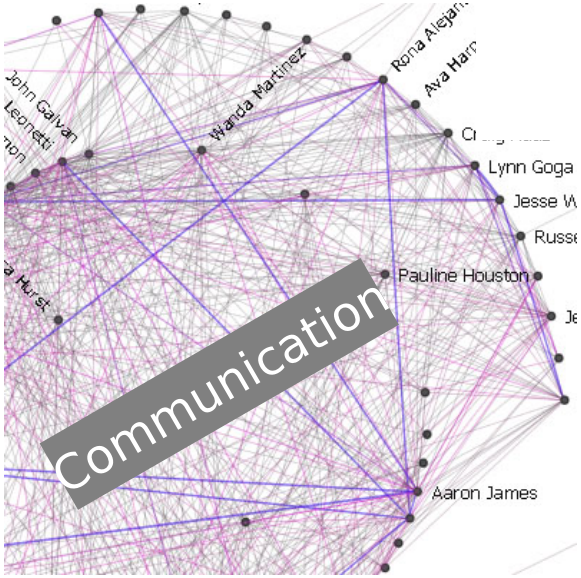
Datum der wissenschaftlichen Aussprache:	9. November 2011
Vorsitz des Promotionsausschusses:	Prof. Dr. Karin Harbusch
Berichterstatter:	Prof. Dr. Steffen Staab
Berichterstatter:	Prof. Dr. Christian Bauckhage
Berichterstatter:	Prof. Dr. Klaus Obermayer

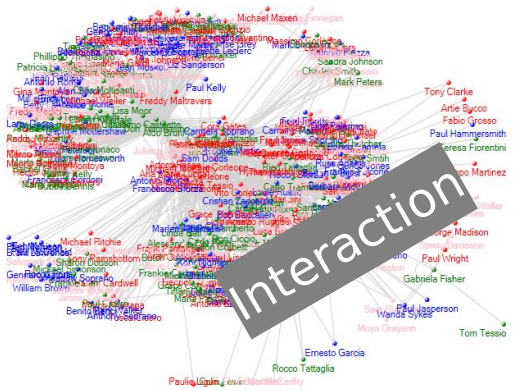

The trick
is...

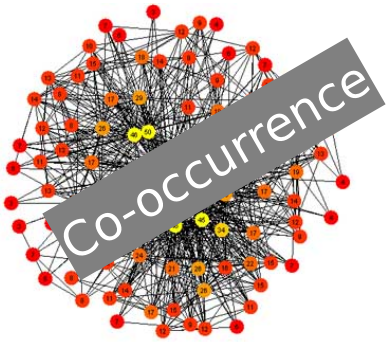

Everything
is a
NETWORK!

Well, Only *Almost* Everything Is a Network

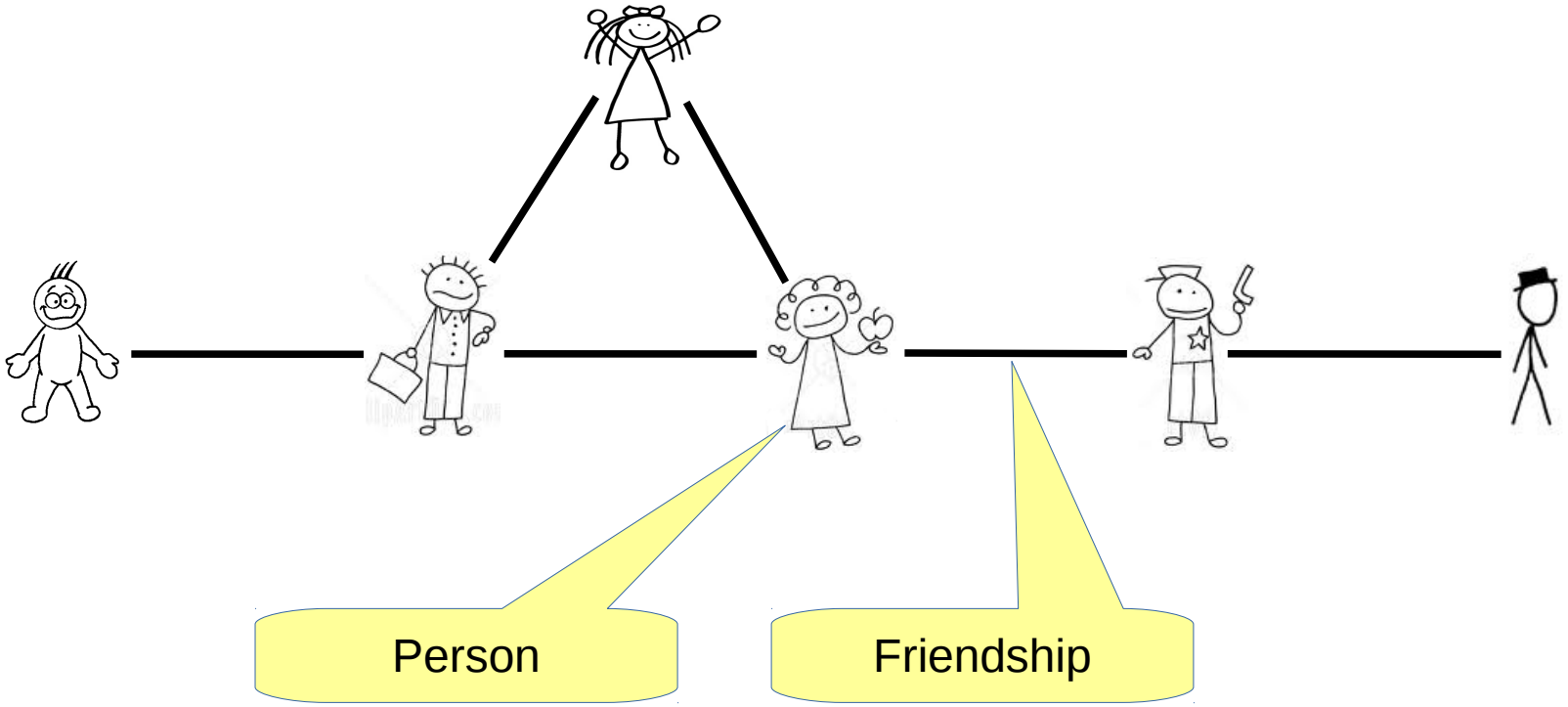




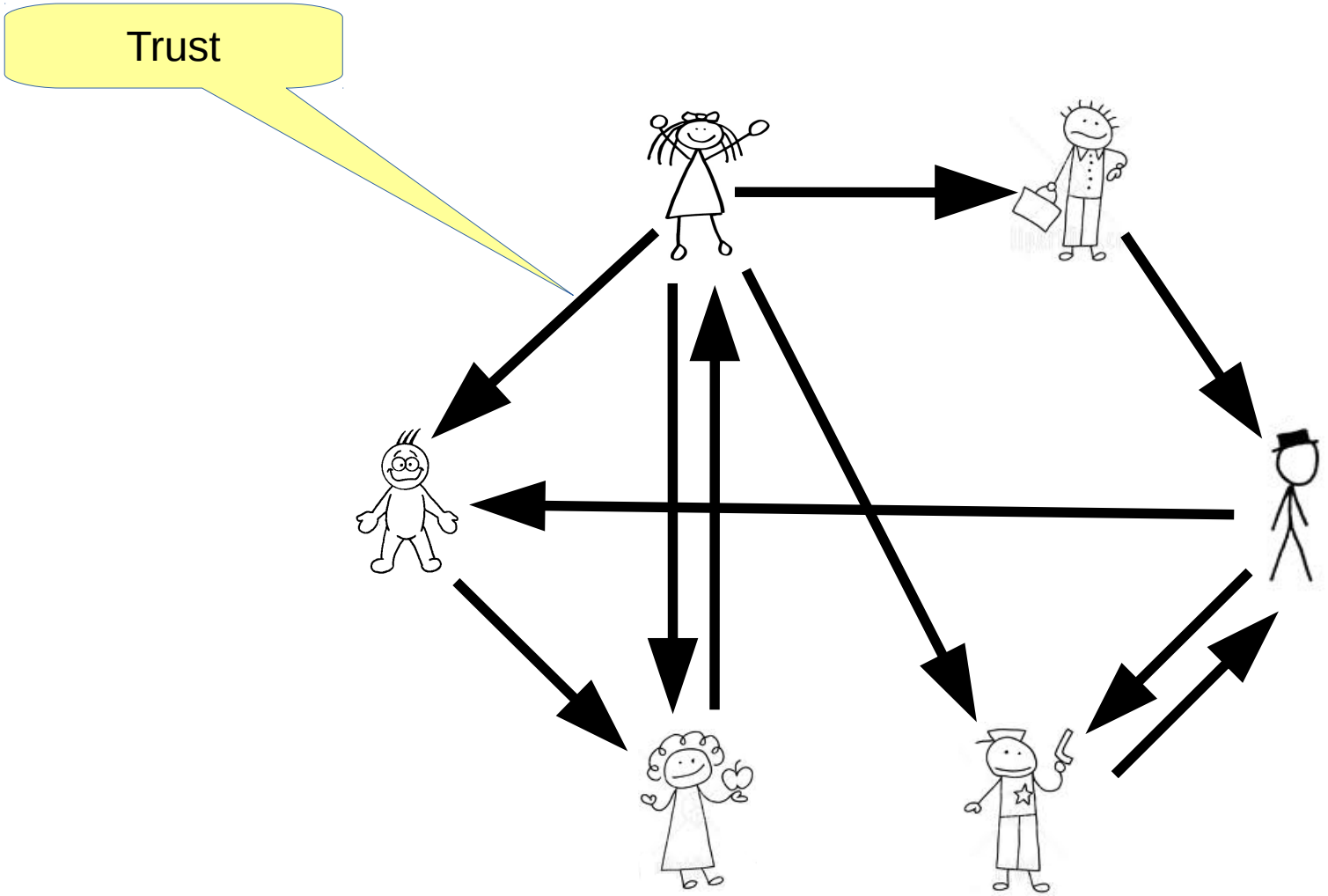




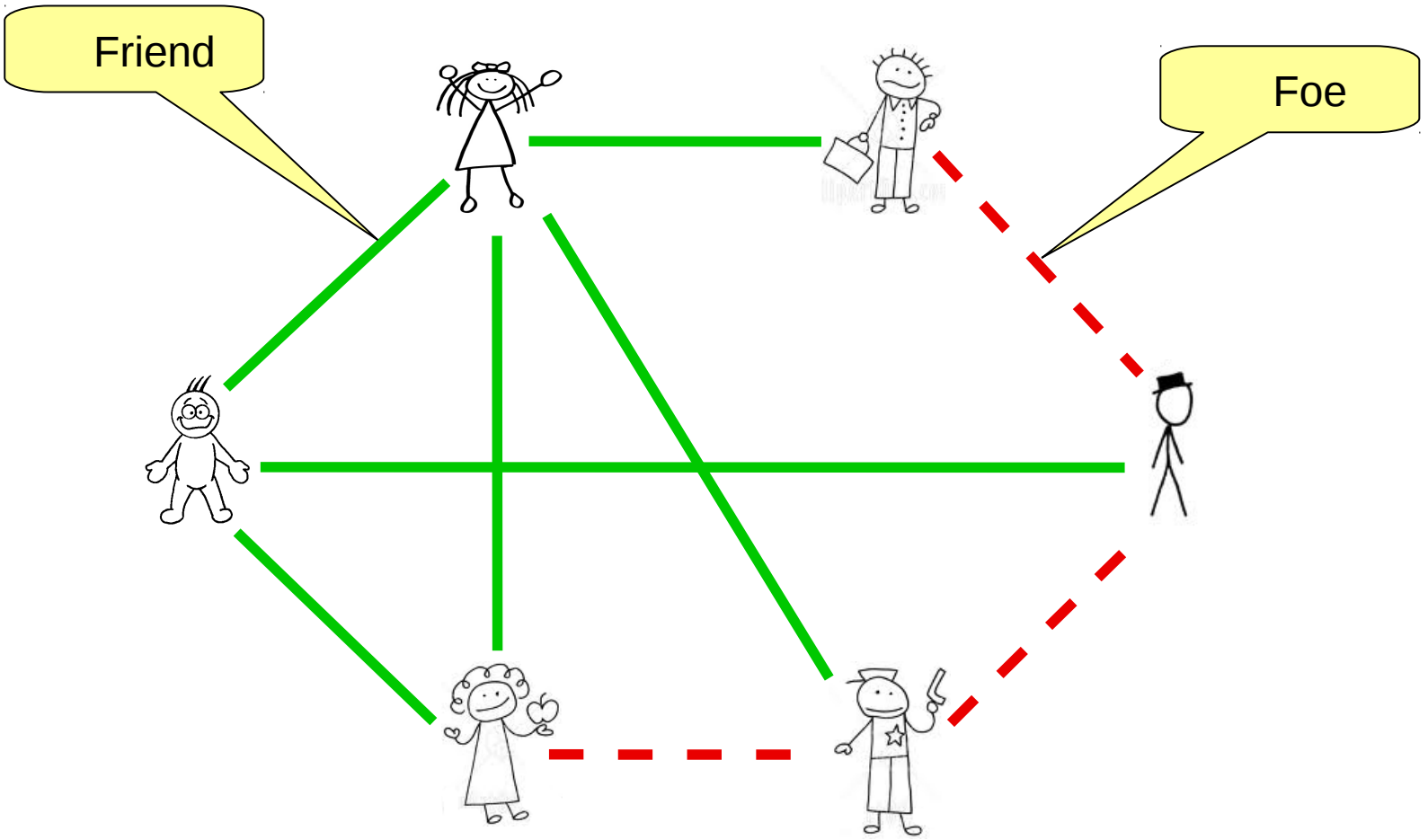
Social Network



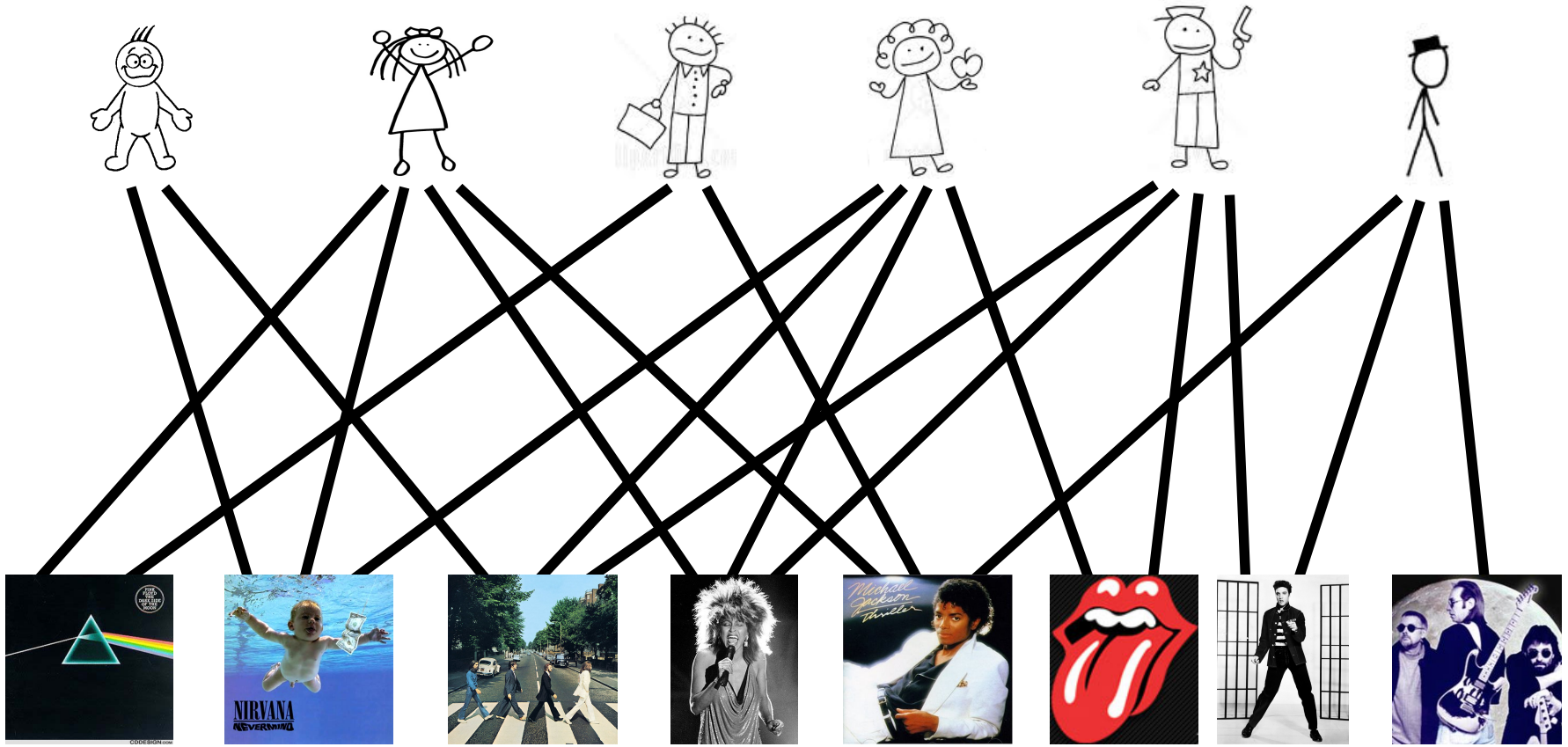
Trust Network



Signed Network



Bipartite Network



A Network Dataset Is Like a Gummi Bear

A Network Dataset Is Like a Gummi Bear

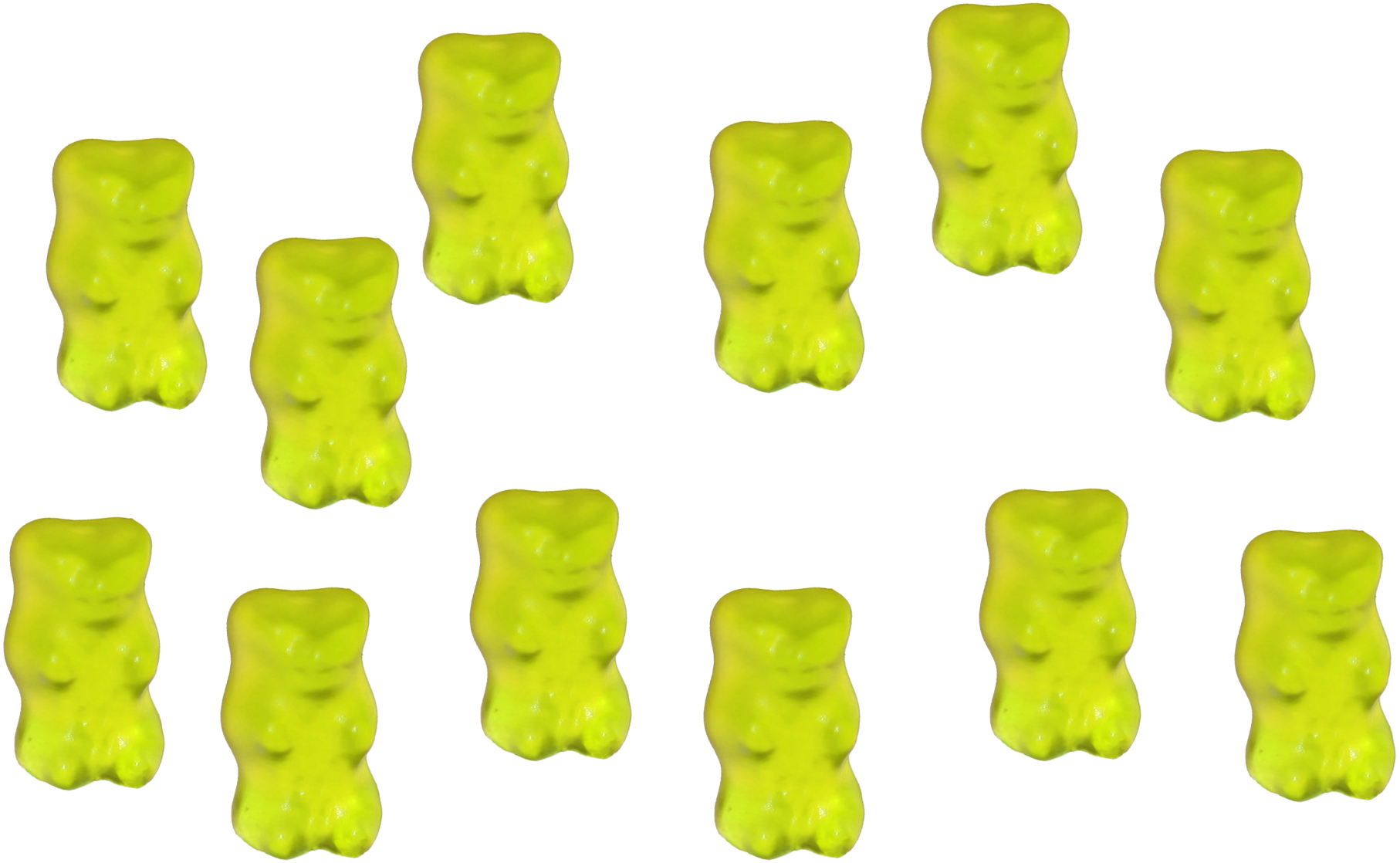
Lots of content to analyse

Test network models

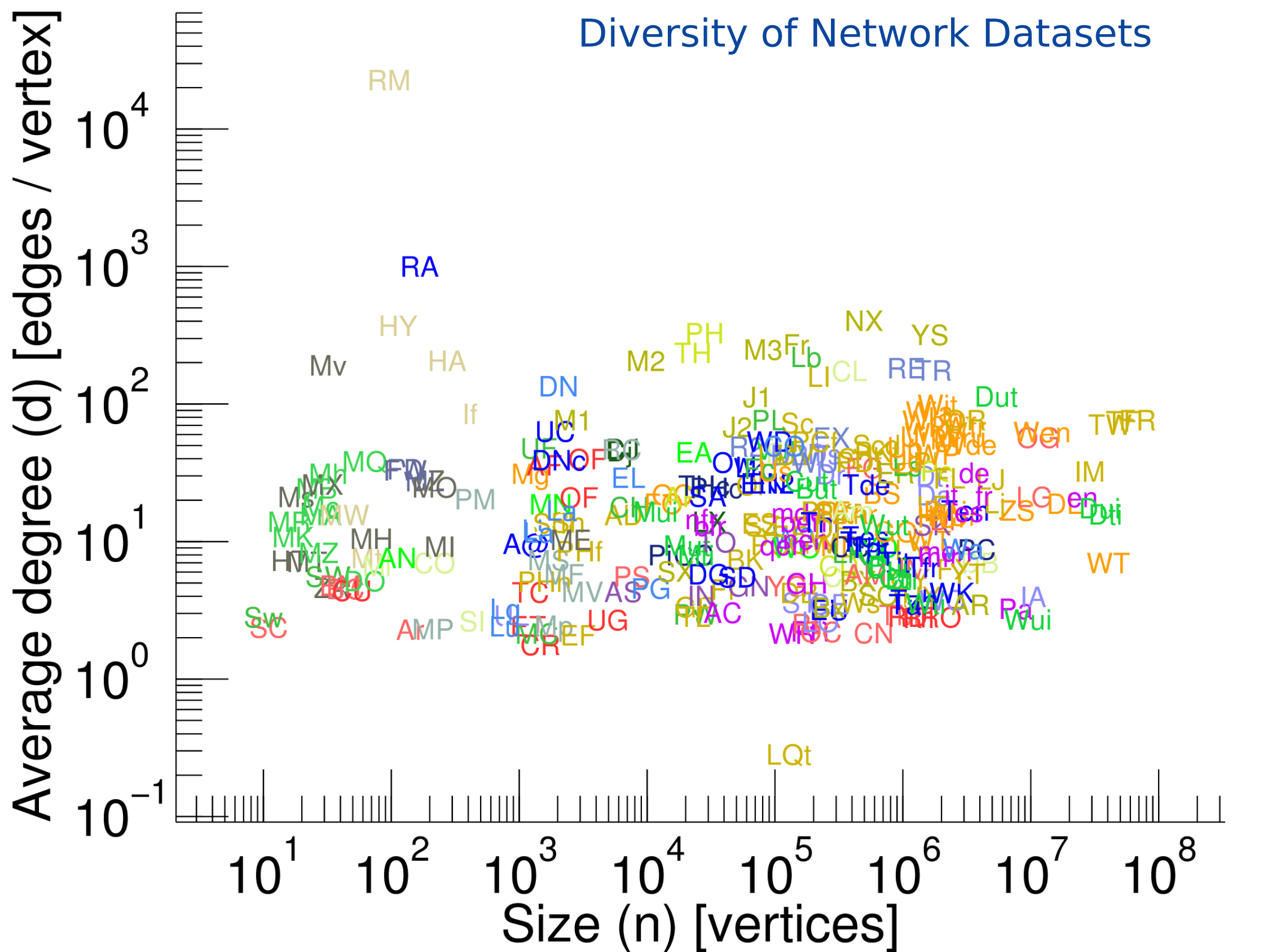
Evaluate prediction algorithms

Test scalability of algorithms

When You Have Tested One, You Have Tested All ?!



Or Do You?

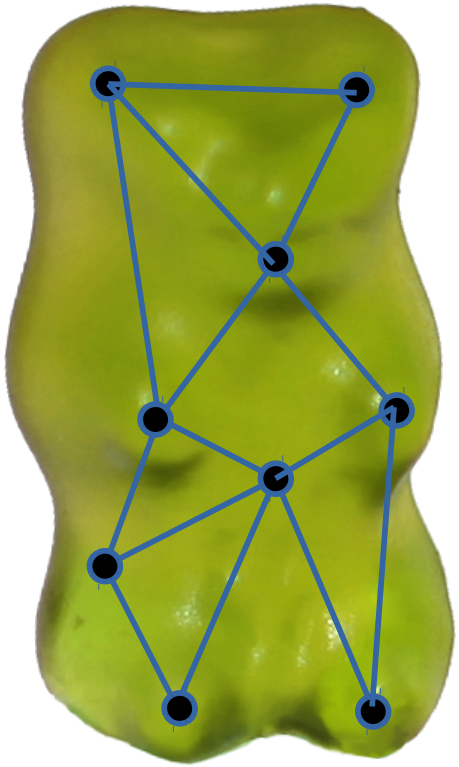

Diversity of Network Datasets
Average degree (d) [edges / vertex]
Size (n) [vertices]
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{1}
10^{2}
10^{3}
10^{4}
10^{5}
10^{6}
10^{7}
10^{8}
RM
RA
HY
Mv
HA
If
DN
M1
M2
TH
PH
M3
Fr
Lb
NX
YS
J1
J2
PLSc
Dut
TWFR
IM
LQt

Network Categories

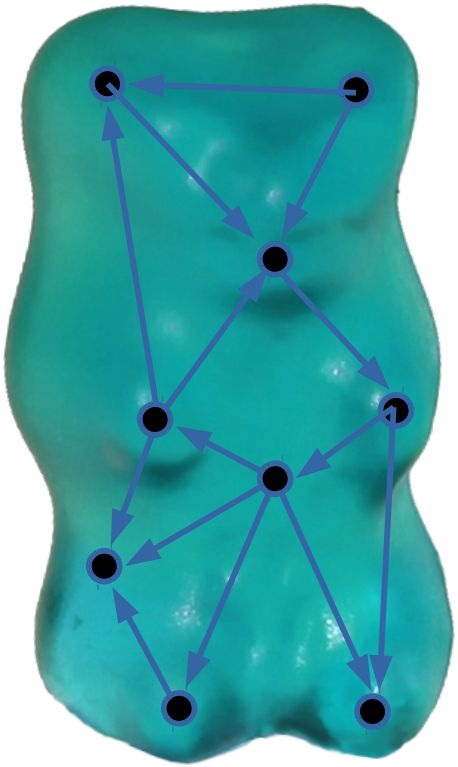
	Category	Vertices	Edges	Properties	Count
●	Affiliation	Actors, groups	Membership	B − =	16
●	Animal	Animals	Tie	U D − +	9
●	Authorship	Authors, works	Authorship	B − =	18
●	Citation	Documents	Citation	D −	6
●	Coauthorship	Authors	Coauthorship	U − =	5
●	Communication	Persons	Message	U D − =	24
●	Computer	Computers	Connection	U D − =	5
●	Feature	Items, features	Property	B − = +	11
●	Folksonomy	Users, tags, items	Tag assignment	B =	18
●	HumanContact	Persons	Real-life contact	U = +	6
●	HumanSocial	Persons	Real-life tie	U D − + ± ≛	11
●	Hyperlink	Web page	Hyperlink	D B − = ⇄	30
●	Infrastructure	Location	Connection	U D − = +	11
●	Interaction	Persons, items	Interaction	D B − = ≛	11
●	Lexical	Words	Lexical relationship	U D − =	5
●	Metabolic	Metabolites	Interaction	U D − =	7
●	Misc	Various	Various	U D − = + ≛	8
●	OnlineContact	Users	Online interaction	U D − = ± ≛ ⇄	11
●	Rating	Users, items	Rating	B − = * $*^*$	15
●	Social	Persons	Online tie	U D − + ± * ⇄	45
●	Software	Software Component	Dependency	D − =	3
●	Text	Documents, words	Occurrence	B =	5
●	Trophic	Species	Carbon exchange	D − +	3
	Total				283

Network Formats

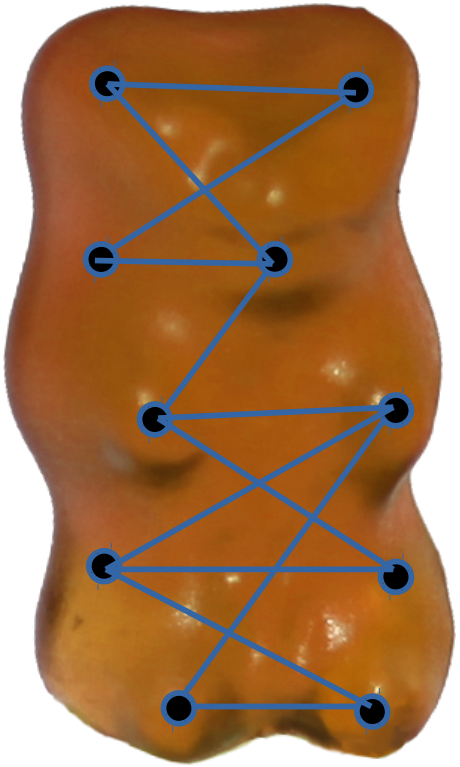
U • Undirected



D • Directed



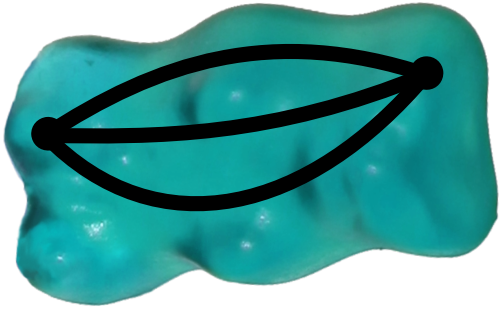
B • Bipartite



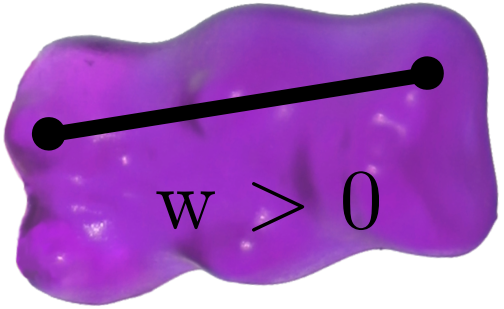
Edge Weight and Multiplicity Types

– • Unweighted

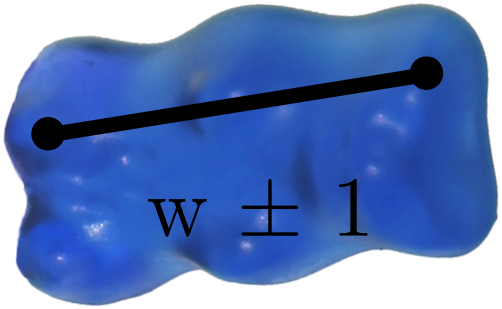
= • Multiple



+ • Positive

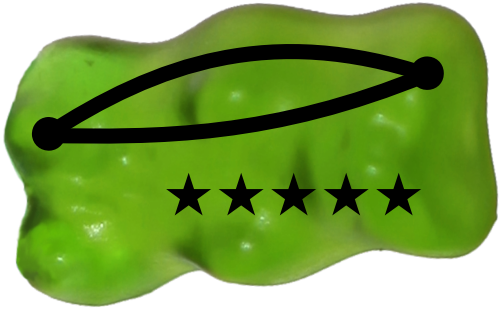


± • Signed

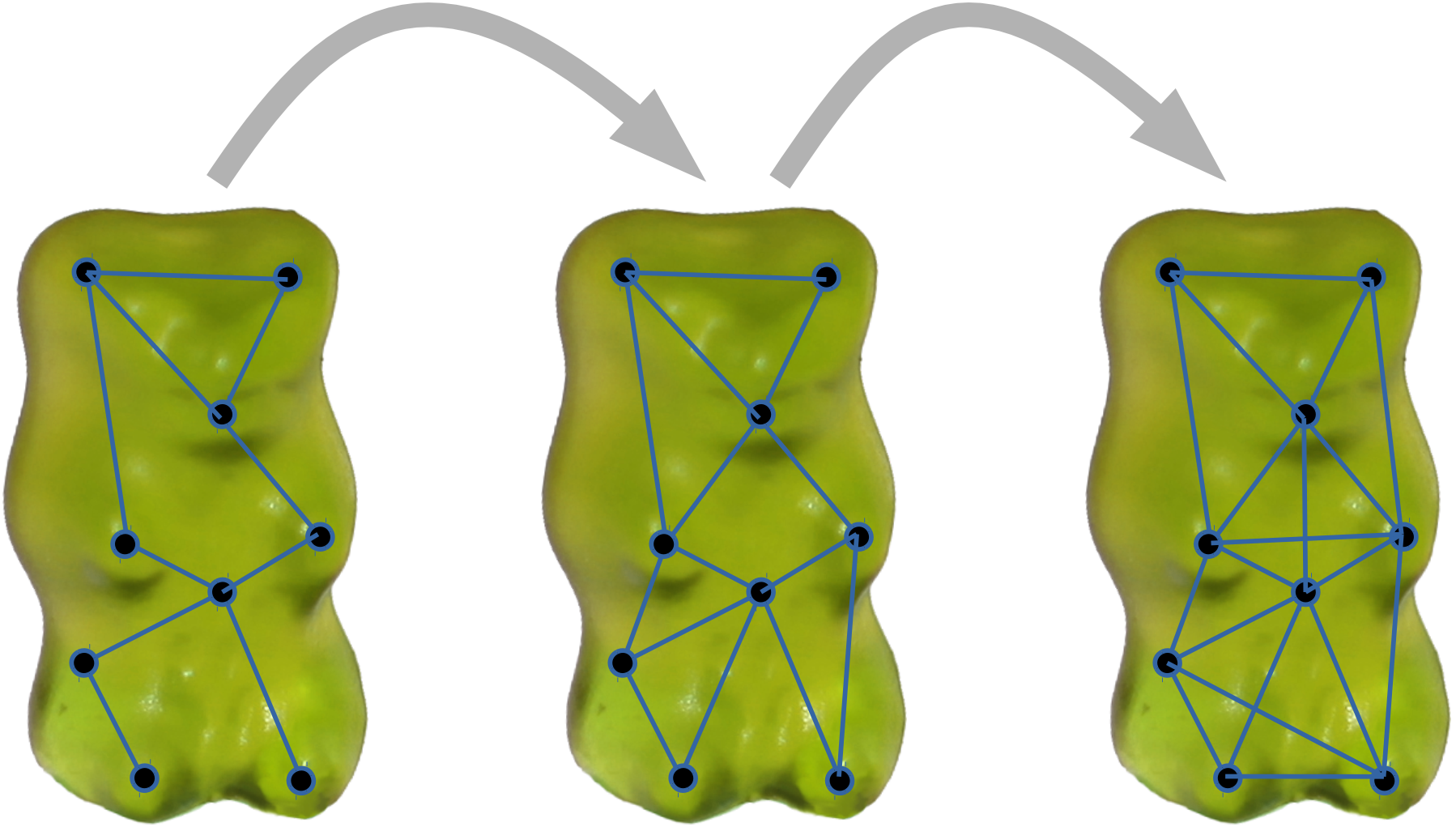


* • Rating

$*^{*}$ • Multiple Ratings



Timestamps



Example Dataset

http://konect.uni-koblenz.de/networks/slashdot-zoo

KONECT

Home Networks Statistics Plots Downloads Toolbox Publications Help License About

KONECT > Networks > **Slashdot Zoo**

Slashdot Zoo

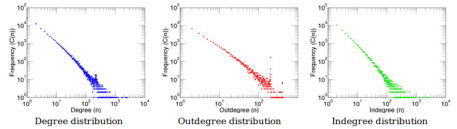
About this network

This is the signed social network of users of the technology news site Slashdot (slashdot.org), connected by directed "friend" and "foe" relations. The "friend" and "foe" labels are used on Slashdot to mark users, and influence the scores as seen by each user. For instance, If user A marks user B as a foe, the score of user B's posts will be decreased as shown to user A.

Network info

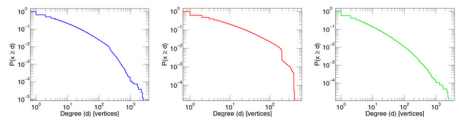
Code	**SZ**
Category	Social
Data source	http://dai-labor.de/IRML/datasets
Description	User–user friend/foe
Format	Directed
Edge weights	Signed
Size	79,120 vertices
Volume	515,581 edges
Average degree	13.03 edges / vertex
Maximum degree	2,546 edges
Fill	8.236×10^{-5} edges / vertex^2
Largest connected component	79,120 vertices (Network is connected)
Largest strongly connected component	26,997 vertices
Clustering coefficient	2.376%
Relative signed clustering coefficient	0.72513
Algebraic connectivity	0.0095167
Spectral norm	52.237
Effective diameter	5.2187 edges (estimated)
Gini coefficient	77.426%
Power law exponent (estimated) with d_{min}	2.181 (d_{min} = 31)
Relative edge distribution entropy	87.087%
Reciprocity	18.505%

Degree distributions



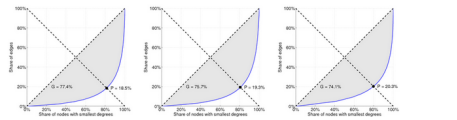
Degree distribution Outdegree distribution Indegree distribution

Cumulative degree distributions



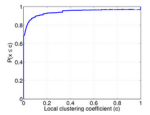
Degree distribution Outdegree distribution Indegree distribution

Lorenz curves



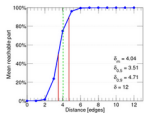
Degree distribution Outdegree distribution Indegree distribution

Clustering coefficient distribution



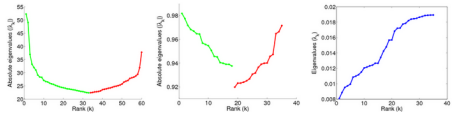
Clustering coefficient distribution

Hop plot



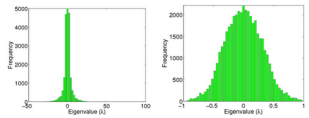
Hop plot

Top-*k* eigenvalues



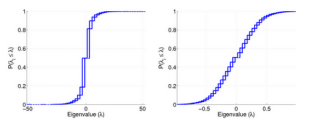
Top-*k* eigenvalues of **A** Top-*k* eigenvalues of **N** Top-*k* eigenvalues of **L**

Spectral distributions



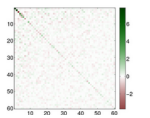
Eigenvalues of **A** Eigenvalues of **N**

Cumulative spectral distributions



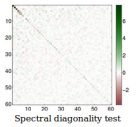
Eigenvalues of **A** Eigenvalues of **N**

Spectral diagonality test



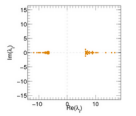
Spectral diagonality test

Spectral diagonality test



Spectral diagonality test

Complex eigenvalues of the asymmetric adjacency matrix



Complex eigenvalues of the asymmetric adjacency matrix

Downloads

TSV file:	slashdot-zoo.tar.bz2 (1.88 MiB)
Extraction code:	zoo.tar.bz2 (26.14 KiB)
Full dataset:	slashdot-zoo.tar.bz2 (18.05 MiB)

References

[1] Slashdot zoo network dataset - KONECT, October 2013. [http]

[2] Jérôme Kunegis, Andreas Lommatzsch, and Christian Bauckhage. The Slashdot Zoo: Mining a social network with negative edges. In *Proc. Int. World Wide Web Conf.*, pages 741-750, 2009. [.pdf]

BibTeX

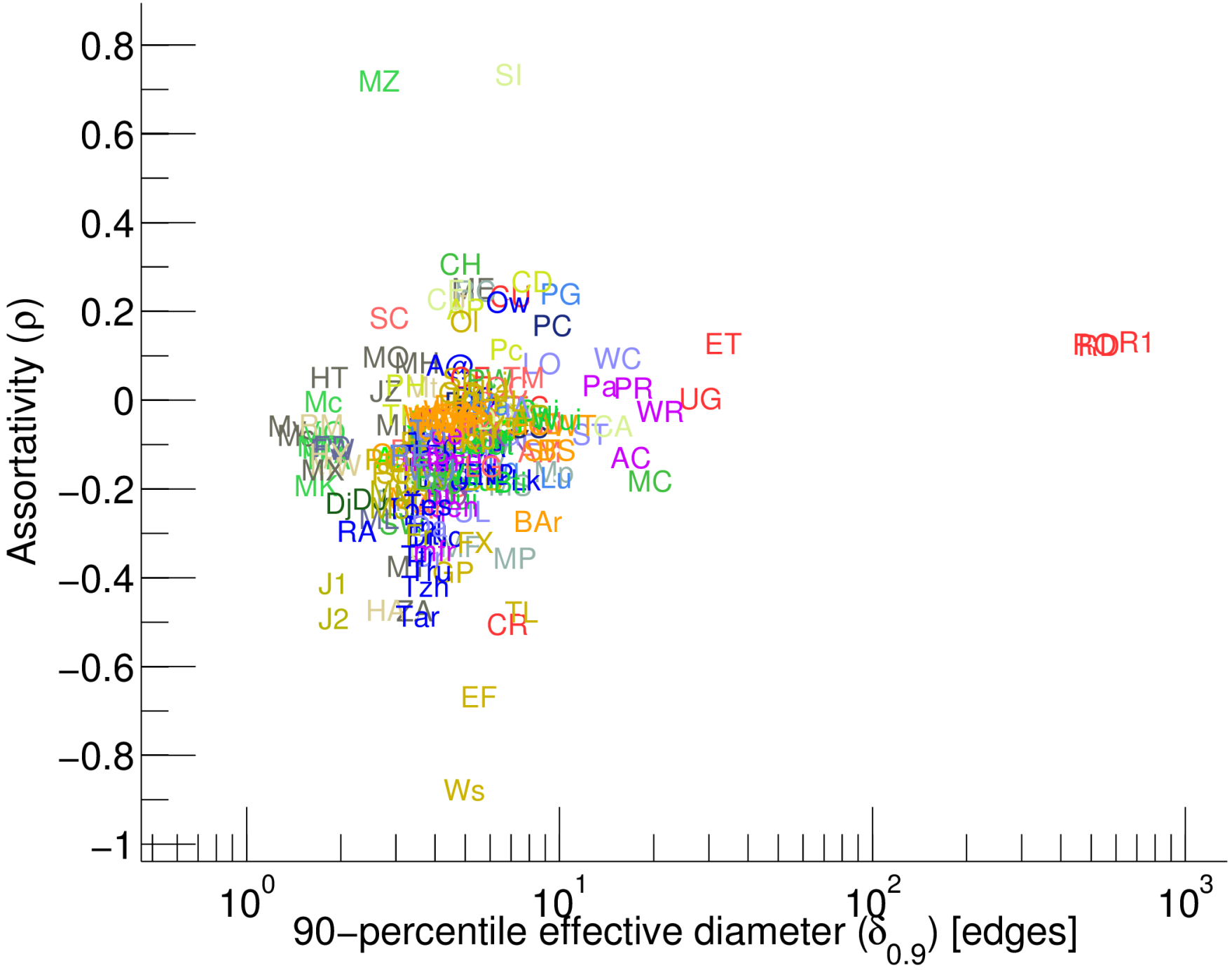
Updated on Wed Oct 16 13:03:08 2013 | Privacy by Jérôme Kunegis

UNIVERSITÄT KOBLENZ · LANDAU WeST People and Knowledge Networks ROBUST GPLv3

Network Comparison

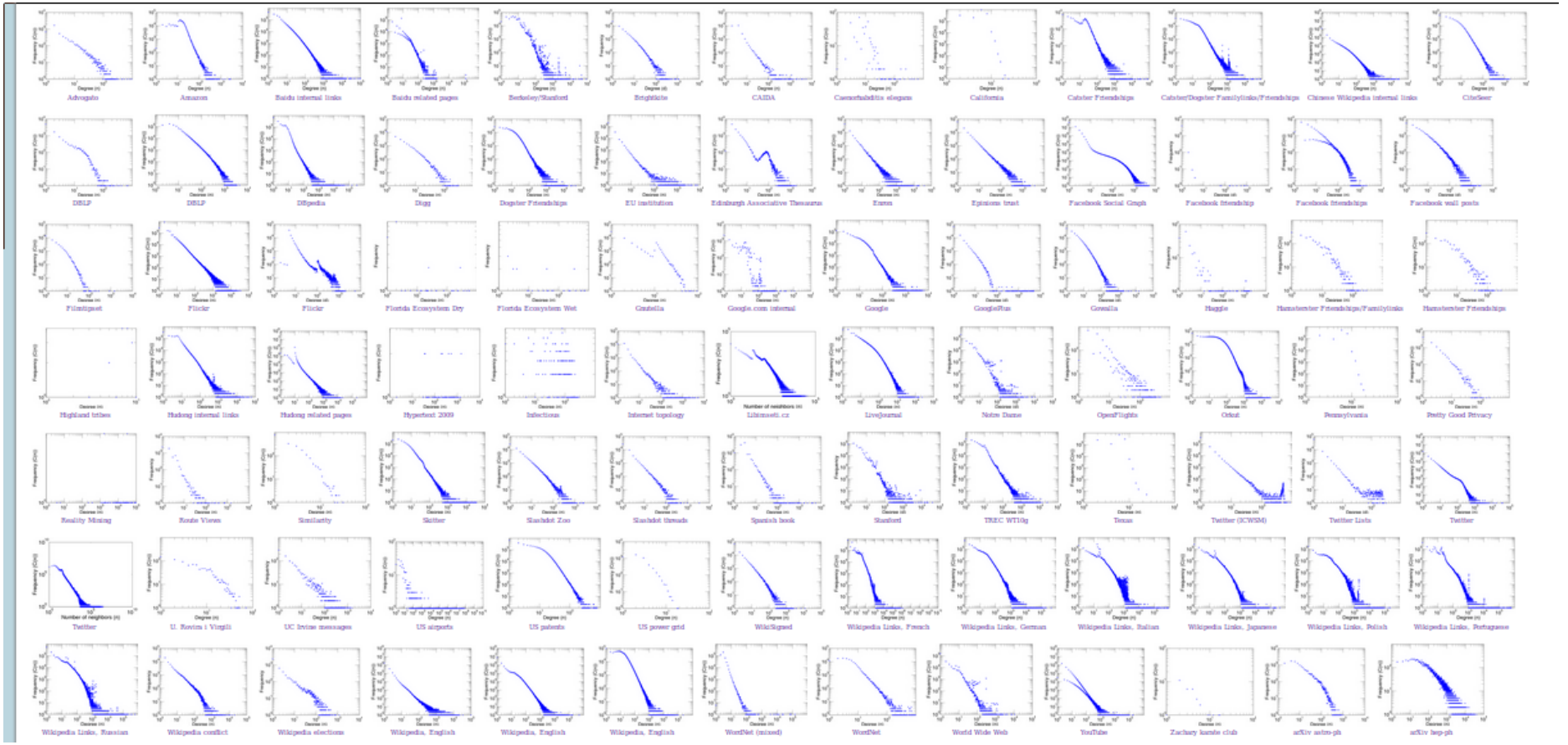
http://konect.uni-koblenz.de/statistics/

Command: `stu @scatter.diameff90.assortativity`



Network Comparison: Plots

http://konect.uni-koblenz.de/plots/

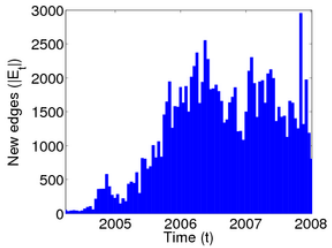


Example: Degree distribution

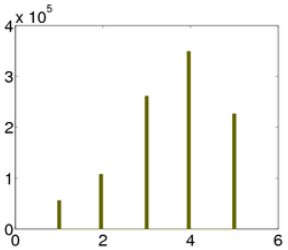
Command: `stu @degree`

More Plots

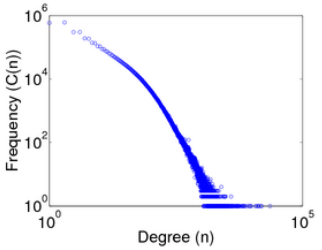
http://konect.uni-koblenz.de/plots/



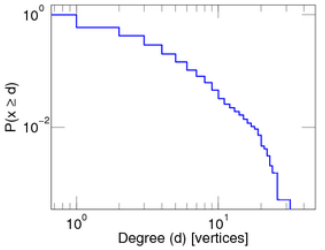
Temporal distribution



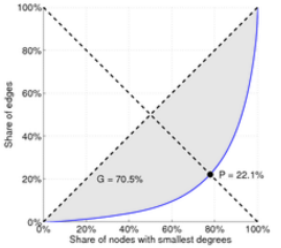
Edge weight and multiplicity distribution



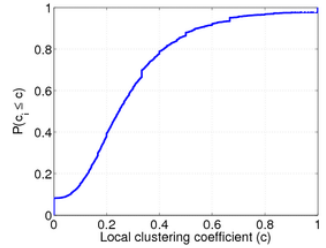
Degree distribution



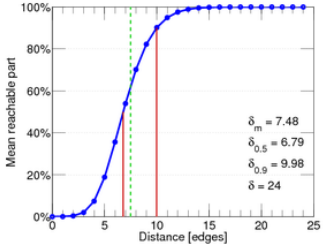
Cumulative degree distribution



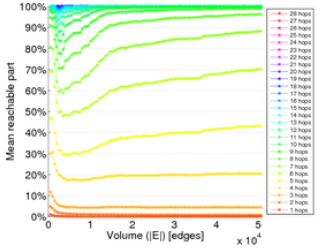
Lorenz curve



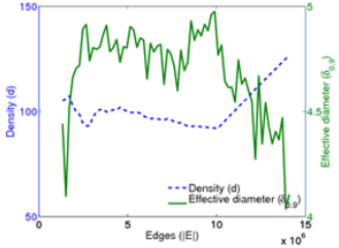
Clustering coefficient distribution



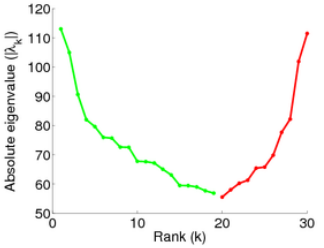
Hop plot



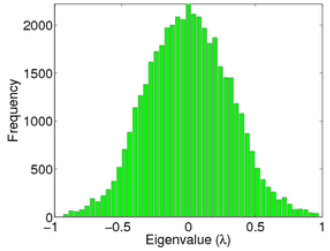
Temporal hop plot



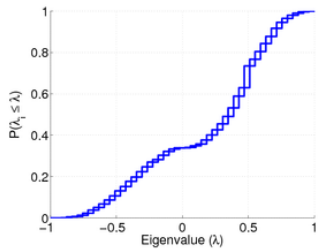
Evolution of average degree and diameter



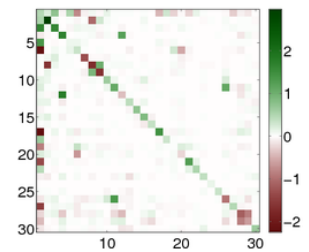
Top-*k* eigenvalues



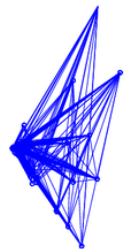
Spectral distribution



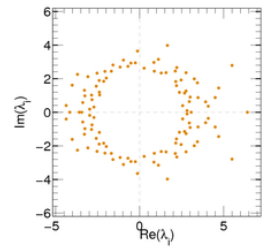
Cumulative spectral distribution



Spectral diagonality test



Network visualization



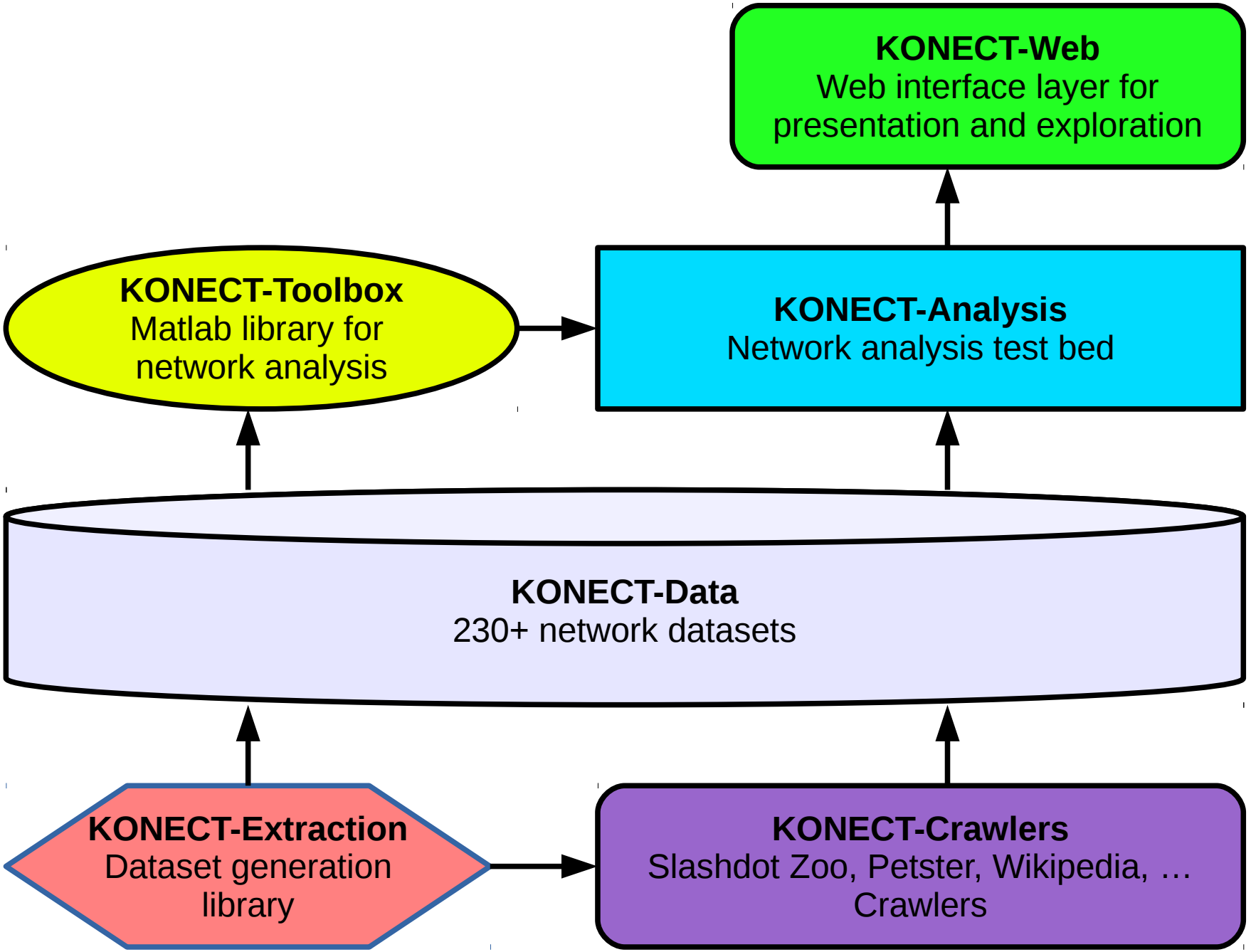
Complex eigenvalues of the asymmetric adjacency matrix

Download

http://konect.uni-koblenz.de/networks/

Code	Name	Category	F.	W.	M.	Size	Volume	Avg. degree	Download
Ds	discogs_lstyle	Features	B	▬		244,147	5,255,950	21.56	
AN	Adjective-noun relationships	Lexical	B	−		194	425	3.94	
AD	Advogato	Social	D	+		6,551	51,332	7.84	
AM	Amazon	Contact	U	−		805,751	3,387,388	4.20	
AR	Amazon ratings	Ratings	B	*	⊘	3,376,972	5,838,041	2.72	
AP	arXiv astro-ph	Contact	U	−		37,544	396,160	10.55	
AC	arXiv cond-mat	Authorship	B	−		38,741	58,595	3.50	
PH	arXiv hep-ph	Contact	U	−		28,093	12,730,096	453.14	
PHc	arXiv hep-ph	Reference	D	−		60,366	421,578	6.98	
THc	arXiv hep-th	Reference	D	−		48,239	352,807	7.31	
TH	arXiv hep-th	Contact	U	−		22,908	11,209,368	489.32	
th	arXiv hep-th (KDD Cup)	Reference	D	−		27,770	352,807	12.70	
BAI	Baidu	Reference	D	▬		2,141,300	17,794,839	8.31	
BAr	Baidu	Reference	D	▬		415,641	3,284,387	7.90	
BS	Berkeley/Stanford	Reference	D	−		1,297,360	7,600,595	5.86	
Btl	BibSonomy ti	Folksonomy	B	▬	⊘	975,963	2,555,080	12.48	
Bul	BibSonomy ui	Folksonomy	B	▬	⊘	777,064	2,555,080	440.99	
But	BibSonomy ut	Folksonomy	B	▬	⊘	210,467	2,555,080	440.99	
BK	Brightkite	Social	U	−		58,228	214,078	3.68	
PM	Caenorhabditis elegans	Contact	U	▬		453	4,596	10.15	
IN	CAIDA	Physical	U	−		26,475	106,762	4.03	
RO	California	Physical	U	−		3,930,412	5,533,214	1.41	
Sc	Catster	Social	U	−		149,700	5,449,275	36.40	
Scd	Catster/Dogster	Social	U	−		624,127	15,705,337	25.16	
CS	CiteSeer	Reference	D	−		723,131	1,764,929	2.44	
Ctl	CiteULike ti	Folksonomy	B	▬	⊘	885,046	2,411,819	15.74	
Cul	CiteULike ui	Folksonomy	B	▬	⊘	754,484	2,411,819	106.18	
Cut	CiteULike ut	Folksonomy	B	▬	⊘	175,992	2,411,819	106.18	
CN	Countries	Affiliation	B	▬		512,761	557,587	1.09	
PI	DBLP	Reference	D	−		12,591	49,793	3.95	

KONECT Project Overview



Matlab Toolbox

https://github.com/kunegis/konect-toolbox

```
konect_dentropy.m
konect_diameff.m
konect_diammean.m
konect_effective_diameter.m
konect_eigl.m
konect_eign.m
konect_eigskew.m
konect_first_index.m
konect_fromto.m
konect_gini_direct.m
konect_gini.m
konect_hopdistr_ex.m
konect_hopdistr.m
konect_imageubu_complex.m
konect_imageubu.m
konect_jain.m
konect_join.m
konect_label_statistic.m
konect_map.m
konect_matrix.m
konect_mauc.m
konect_network_rank_abs.m
konect_normalize_additively.m
konect_normalized_entropy.m
konect_normalize_matrix.m
konect_normalize_rows.m
konect_order_dedicom.m
konect_own.m
```

Handbook of Network Analysis

Handbook of Network Analysis

KONECT project

Jérôme Kunegis
University of Namur, Belgium
naXys – Namur Center for Complex Systems
with web hosting provided by the Institute of Web Science and Technologies (WeST)
at the University of Koblenz–Landau, Germany
konect.uni-koblenz.de

June 26, 2017

Abstract

This is the handbook for the KONECT project, a scientific project to archive network datasets, compute systematic network theoretic statistics about them, visualize their properties, and provide corresponding data and Free Software

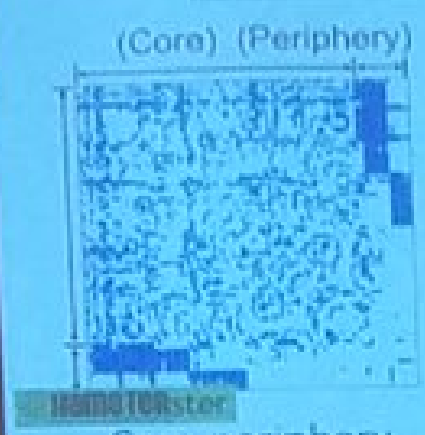
https://github.com/kunegis/konect-handbook

methods, definitions and conventions used in the project, and serves as a general handbook of network mining, with an emphasis on spectral graph theoretical methods, i.e., such methods that are based on the use of specific characteristic matrices of graphs.

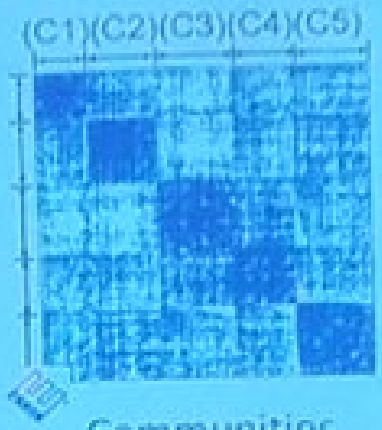
KONECT Datasets in the Wild

Structural-Core Pattern: Pattern

Pattern 3: Structural-Core Pattern
Degeneracy-cores have structural patterns such as **core-periphery** and **communities**



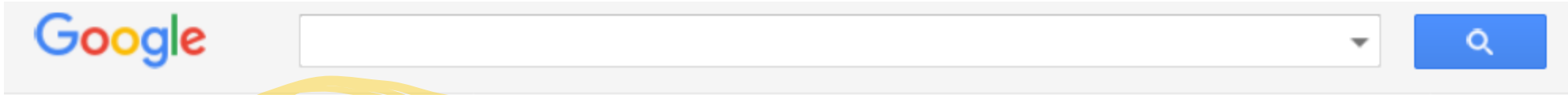
Core-periphery



Communities

Nodes in degeneracy-cores are not homogeneous

(Tina Eliassi-Rad, NetSci 2017)



Scholar

About 203 results (0.02 sec)

All citations

Articles

Case law

My library

Any time

Since 2017

Since 2016

Since 2013

Custom range...

Sort by relevance

Sort by date

include citations

Create alert

Konect: the koblenz network collection

Search within citing articles

Recent advances in graph partitioning

A Buluç, H Meyerhenke, I Safro, P Sanders... - Algorithm ..., 2016 - Springer

Cited by 76 Related articles All 13 versions Cite Save More

Estimating clustering coefficients and size of social networks via random walk

SJ Hardiman, L Katzir - ... of the 22nd international conference on World ..., 2013 - dl.acm.org

Abstract Online social networks have become a major force in today's society and economy. The largest of today's social networks may have hundreds of millions to more than a billion users. Such networks are too large to be downloaded or stored locally, even if terms of use

Cited by 44 Related articles All 10 versions Cite Save More

BFS and coloring-based parallel algorithms for strongly connected components and related problems

GM Slota, S Rajamanickam... - Parallel and Distributed ..., 2014 - ieeexplore.ieee.org

Abstract: Finding the strongly connected components (SCCs) of a directed graph is a fundamental graph-theoretic problem. Tarjan's algorithm is an efficient serial algorithm to find SCCs, but relies on the hard-to-parallelize depth-first search (DFS). We observe that

Cited by 39 Related articles All 13 versions Cite Save More

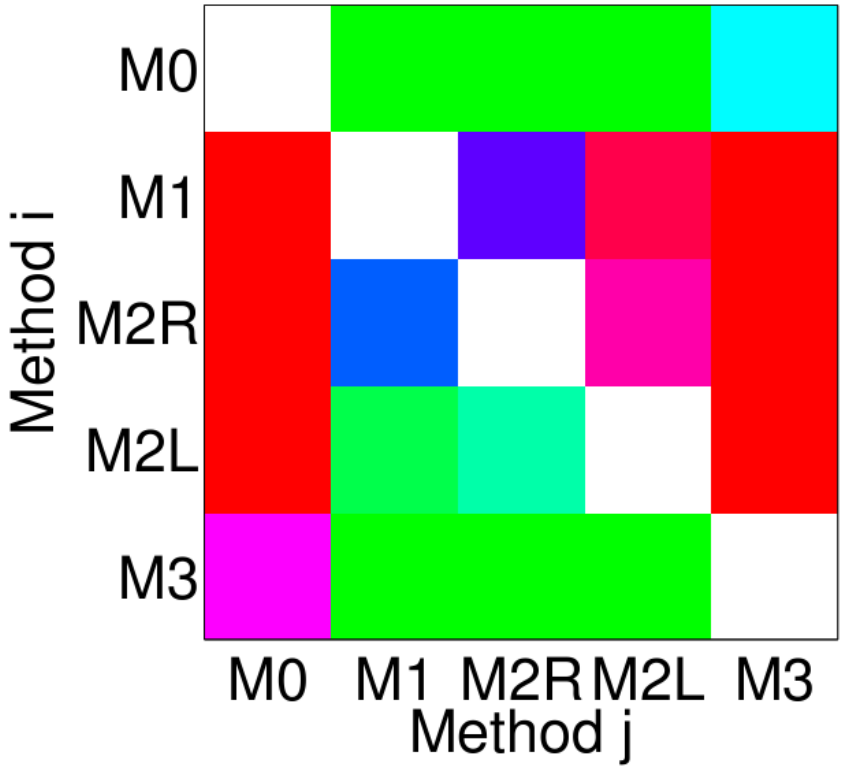
Pulp: Scalable multi-objective multi-constraint partitioning for small-world networks

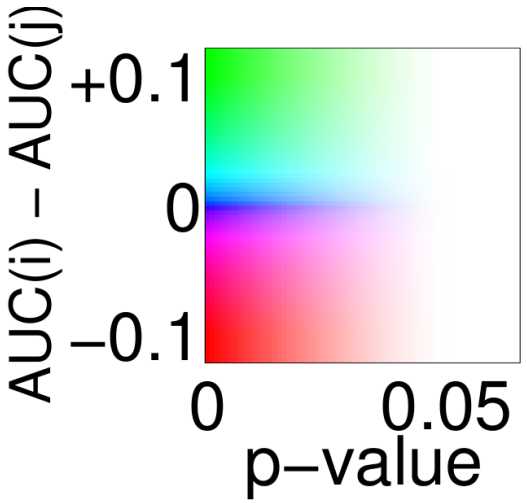
GM Slota, K Madduri... - Big Data (Big Data), 2014 ..., 2014 - ieeexplore.ieee.org

Abstract: We present PuLP, a parallel and memory-efficient graph partitioning method specifically designed to partition low-diameter networks with skewed degree distributions. Graph partitioning is an important Big Data problem because it impacts the execution time

Cited by 19 Related articles All 11 versions Cite Save More

Application 1: Link Prediction



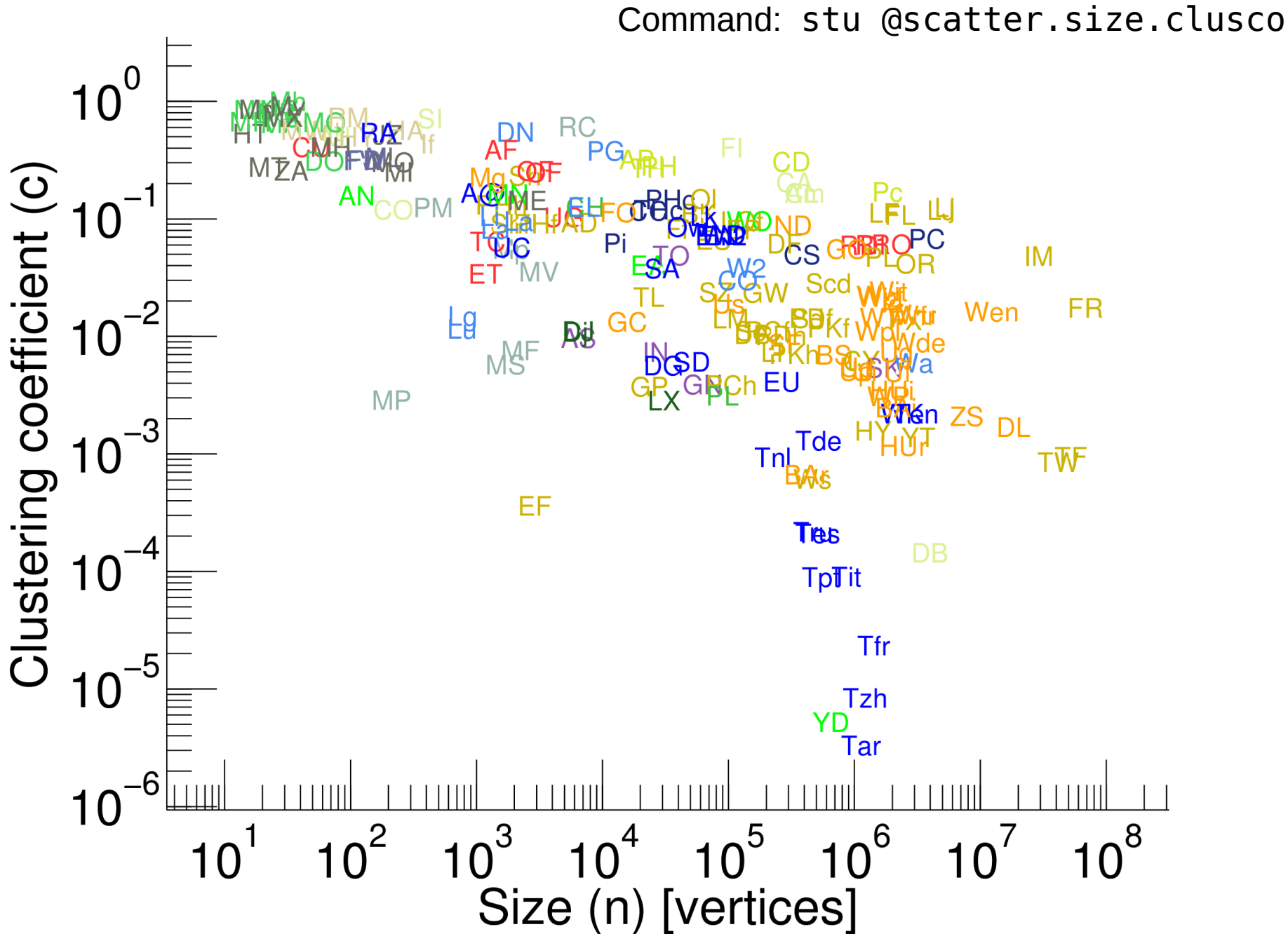


(Preusse & Kunegis, unpublished)

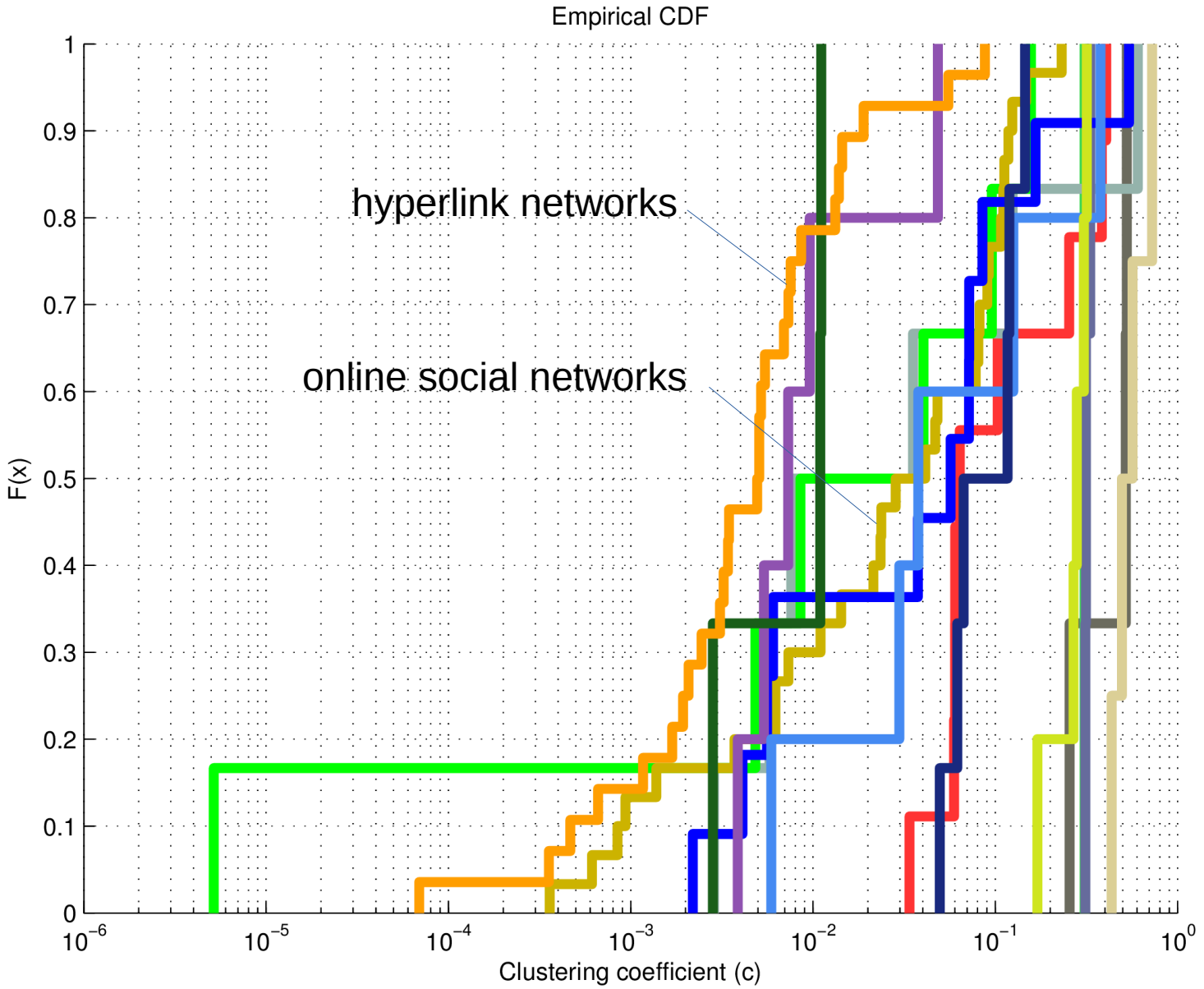
Application 2: Measuring Properties of Networks

- Everyone knows: networks have high clustering coefficient
- What is the typical clustering coefficient of a social network?
- What is the typical clustering coefficient of a hyperlink network?
- How can one answer these types of questions?

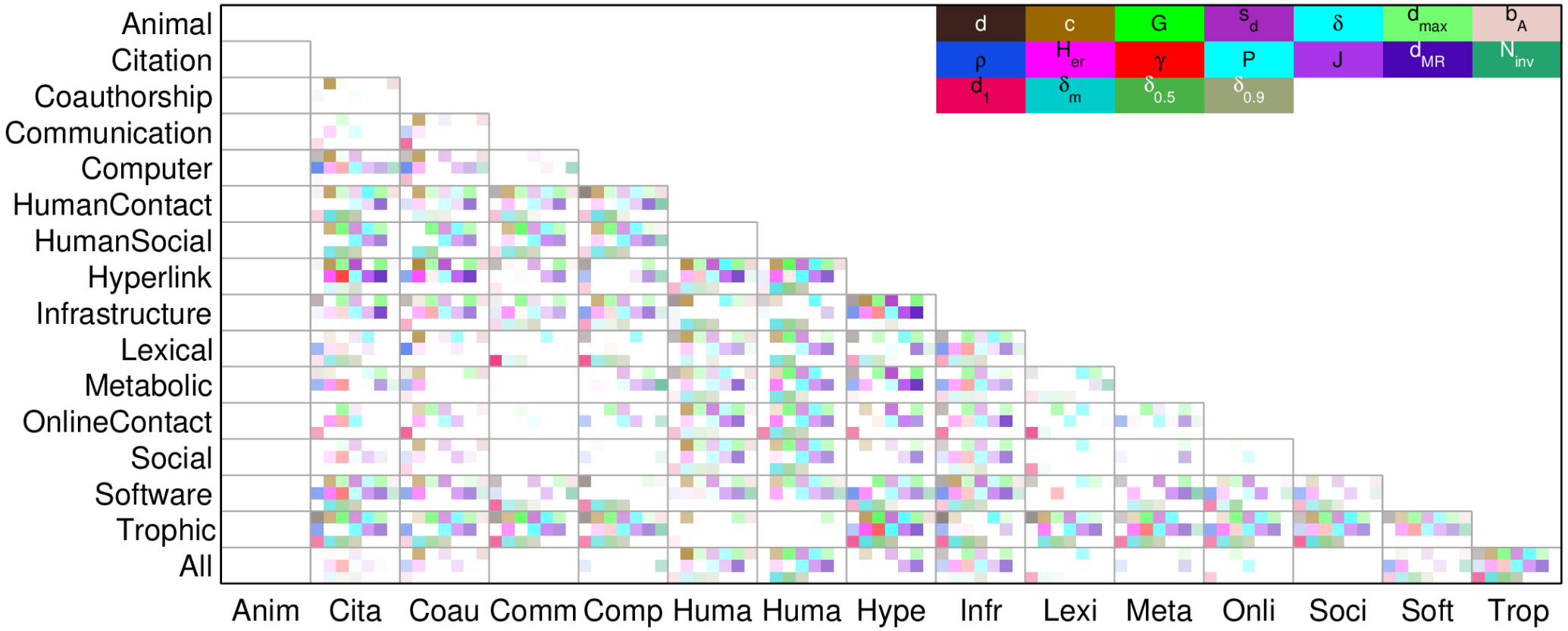
Clustering Coefficient



Can the Category of a Network Be Predicted?

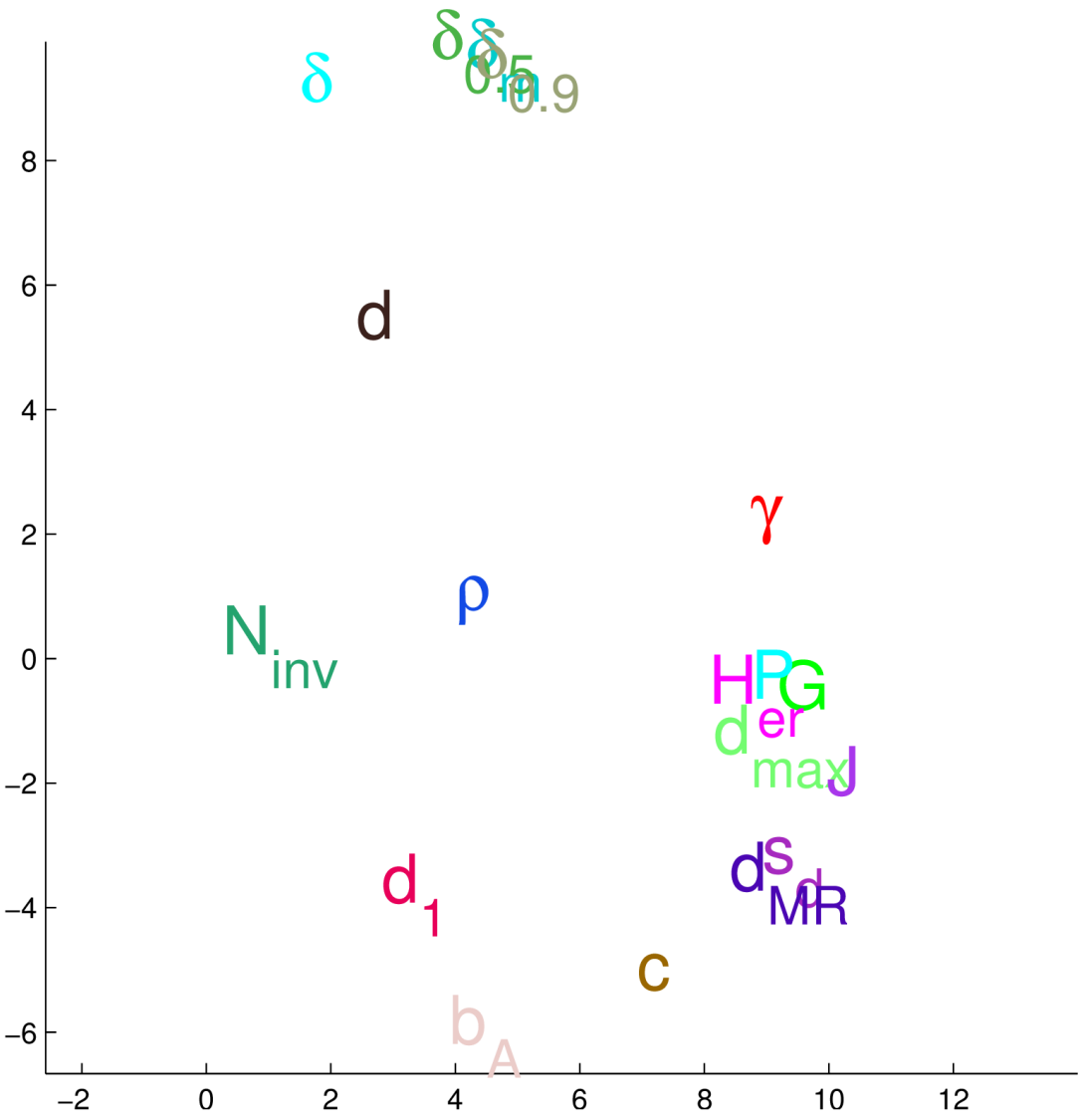


Predicting the Type of a Network



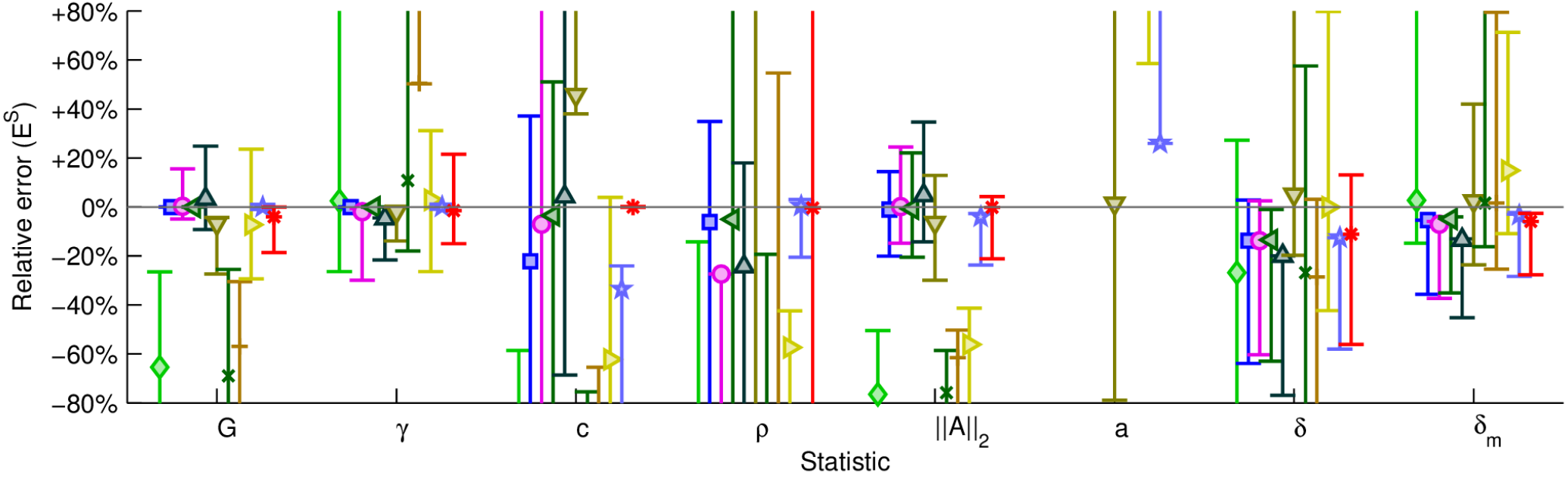
Command: `stu @all-nc`

Clustering of Network Statistics (PCA)



Application 3: Verify Graph Models

- Experiment: Generate graphs with properties matching those of given graphs
- Evaluation: Average relative error on 36 datasets



(Kunegis et al., submitted to Comput. Soc. Netw.)

KONECT Analysis with Stu

- Run tasks in parallel
- Never rebuild a file
- Interrupt at any point

%CPU	Mem [k]	Runtime	Runtime left	Log
40	55068	1-13:49:30		/data/kunegis/tmp/ifub.lasagne-yahoo
41	751920	18:45:53		/data/kunegis/tmp/julia.kunegis.inter2.log.wiki_talk_zh
42	863196	9:54:10		/data/kunegis/tmp/julia.kunegis.inter2.log.wiki_talk_es
43	567820	8:31:25		/data/kunegis/tmp/julia.kunegis.inter2.log.bibsonomy-2ui
43	687976	7:38:18		/data/kunegis/tmp/julia.kunegis.inter2.log.bibsonomy-2ti
42	2094952	5:21:49	173-10:17:18	/data/kunegis/tmp/m.kunegis.hopdistr_time_comp.full.munmun_twitterex_ti.log
43	741520	5:15:42		/data/kunegis/tmp/julia.kunegis.inter2.log.munmun_twitterex_ti
44	1982928	4:26:56	3-23:27:11	/data/kunegis/tmp/m.kunegis.cluscod.youtube-links.log
47	2119532	3:10:17	10-16:11:33	/data/kunegis/tmp/m.kunegis.cluscod.wiki-Talk.log
49	1701260	2:46:27	4-11:24:27	/data/kunegis/tmp/m.kunegis.statistic_comp.squares.wiki-Talk.log
48	1580524	2:23:17	4-17:09:06	/data/kunegis/tmp/m.kunegis.statistic_comp.tour4.wiki-Talk.log
48	2032020	2:11:45	14-07:50:19	/data/kunegis/tmp/m.kunegis.hopdistr_time_comp.full.digg-votes.log
48	905720	2:10:22		/data/kunegis/tmp/julia.kunegis.inter2.log.digg-votes
50	2060064	35:48	6:54:45	/data/kunegis/tmp/m.kunegis.cluscod.petster-cat-friend.log
49	1466944	30:44	2:58:58	/data/kunegis/tmp/m.kunegis.statistic_comp.squares.petster-cat-friend.log
50	1570328	24:46	2:37:59	/data/kunegis/tmp/m.kunegis.statistic_comp.tour4.petster-cat-friend.log

Why Not Use make(1) ?

```
define TEMPLATE_fit

$(foreach NETWORK, $(NETWORKS), fit.$(1).$(NETWORK)): \
fit.$(1).%:  plot/fit.a.$(1).%.eps

dat/fit.$(1).%.mat:  \
  dat/info.% dat/decomposition_split.source.$(1).%.mat \
  dat/split.%.mat dat/means.%.mat m/fit.m
        NETWORK=$$* DECOMPOSITION=$(1) $(OCTAVE) m/fit

plot/fit.a.$(1).%.eps: dat/fit.$(1).%.mat m/fit_plot.m
        NETWORK=$$* DECOMPOSITION=$(1) $(OCTAVE) m/fit_plot

endef
$(foreach DECOMPOSITION, $(DECOMPOSITIONS), $(eval $(call TEMPLATE_fit,$(DECOMPOSITION))))

define TEMPLATE_fit_asym
fit.$(1).all:  $(foreach NETWORK, $(NETWORKS_ASYM), fit.$(1).$(NETWORK))
endef
$(foreach DECOMPOSITION, $(DECOMPOSITIONS_ASYM), $(eval $(call TEMPLATE_fit_asym,$(DECOMPOSITION))))

define TEMPLATE_fit_any
fit.$(1).all:  $(foreach NETWORK, $(NETWORKS), fit.$(1).$(NETWORK))
endef
$(foreach DECOMPOSITION, $(DECOMPOSITIONS_ANY), $(eval $(call TEMPLATE_fit_any,$(DECOMPOSITION))))


$(foreach NETWORK, $(NETWORKS), fit.all.$(NETWORK)): \
fit.all.%:  $(foreach DECOMPOSITION, $(DECOMPOSITIONS_ANY), fit.$(DECOMPOSITION).%)
$(foreach NETWORK, $(NETWORKS_ASYM), fit.all.$(NETWORK)): \
fit.all.%:  $(foreach DECOMPOSITION, $(DECOMPOSITIONS_ASYM), fit.$(DECOMPOSITION).%)
```

Stu Example from KONECT

```
#
# Degree vs local clustering coefficient
#

@degcc:   [dat/dep.degcc];

>dat/dep.degcc:  dat/NETWORKS_SQUARE
{
        for network in $(cat dat/NETWORKS_SQUARE) ; do
                echo @degcc."$network"
        done
}

@degcc.$network:  plot/degcc.a.$network.eps;

plot/degcc.a.$network.eps:
        m/degcc.m $[-t MATLABPATH]
        dat/cluscod.$network.mat
        uni/out.$network
{
        ./matlab m/degcc.m
}
```

Stu 2.5 (In Preparation)

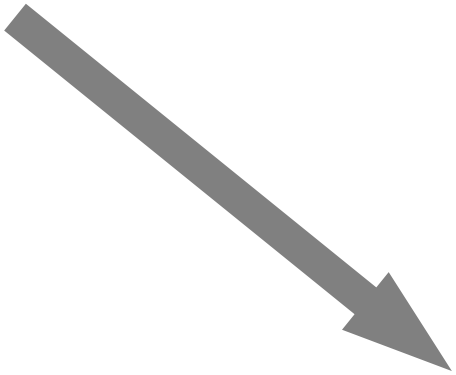
```
@degcc:  @degcc.[dat/NETWORKS_SQUARE];

@degcc.$network:  plot/degcc.$network.eps;

plot/degcc.$network.eps:
   m/degcc.m $[-t MATLABPATH]
    dat/cluscod.$network.mat
   uni/out.$network
{
    ./matlab m/degcc.m
}
```

What Is Stu?

- “What Make should be”
- “A declarative programming language where variables are files”
  - → cf. Unix: “everything is a file”
- “What if the shell was declarative?”
  - → use shell syntax
- Don't hardcode anything
  - → no builtin rules; can be written *in* Stu
- Programming language-agnostic
- Scalability for data mining
  - → Ctrl-C, signals, precious files, O(1), etc.
- “Make” ⇒ “Cook” ⇒ “Stew”
- Free Software, GPLv3
- cf. Stu Feldman, inventor of Make

BONUS: CV with Stu

```
@article{samoilenko:language-hierarchy,
        title = {Linguistic Neighbourhoods: Explaining Cultural
          Borders on {Wikipedia} through Multilingual Co-editing Activity},
        author = {Anna Samoilenko and Fariba Karimi and Daniel Edler
           and Jérôme Kunegis and Markus Strohmaier},
        journal = {Eur. Phys. J. Data Sci.},
        year = {2016},
        volume = {5},
        pages = {9},
        url = {http://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-016-0070-8},
        url_poster = {http://docs.wixstatic.com/ugd/4948de_7ed620c4ddc74bdbb48e3fdd8cb12a9a.pdf},
        url_web = {http://annsamoilenko.wixsite.com/homepage/linguistic-neighbourhoods},
        url_data = {http://annsamoilenko.wixsite.com/homepage/data},
        kunegiscat = {journal},
}
```



[2] Continuous-Time Quantum Walks on Directed Bipartite Graphs. Beat Tödtli, Monika Laner, Jouri Semenov, Beatrice Paoli, Marcel Blattner, Jérôme Kunegis, *Phys. Rev. A*, 94(5):052338, 2016.

[3] Linguistic Neighbourhoods: Explaining Cultural Borders on Wikipedia through Multilingual Co-editing Activity. Anna Samoilenko, Fariba Karimi, Daniel Edler, Jérôme Kunegis, Markus Strohmaier, *Eur. Phys. J. Data Sci.*, 5:9, 2016.

[4] Glaubwürdigkeit und Vertrauen von Online-News. Ines C. Vogel, Jutta Milde, Karin Stengel, Steffen Staab, Christoph Carl Kling, Jérôme Kunegis, *Datenschutz und Datensicherheit*, 40(5):312–316, 2015.

Thank You

Lessons:

- Everything is a network
- Give Jérôme datasets
- Automatize everything

http://konect.uni-koblenz.de/

→ Network contributions accepted ←

Stu: https://github.com/kunegis/stu

KONECT: https://github.com/kunegis/
konect-{handbook,extr,analysis,toolbox}

News via Twitter: @KONECTproject