rocha@lanl.gov

The Active Recommendation Project and MyLibrary

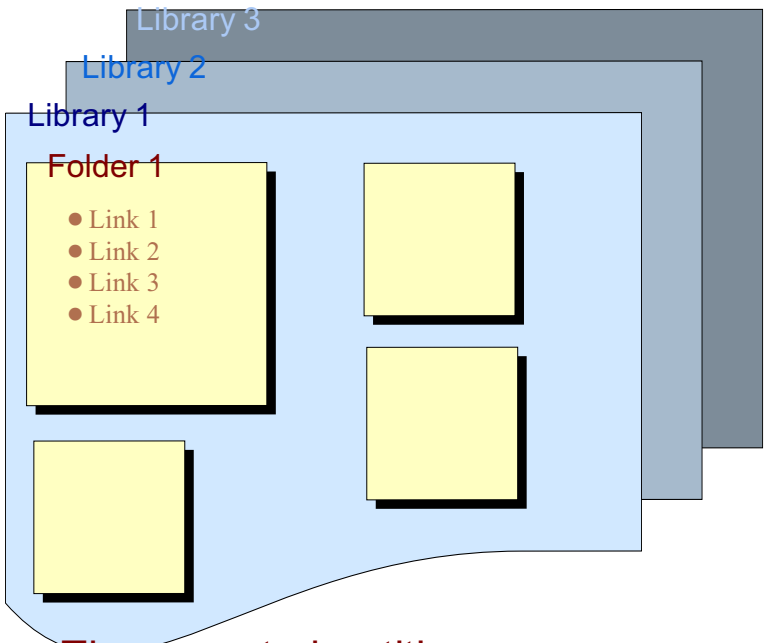
http://arp.lanl.gov

Luis M. Rocha, Andreas Rechtsteiner, Tiago Simas, Chien-Feng Huang, Judith Cohn

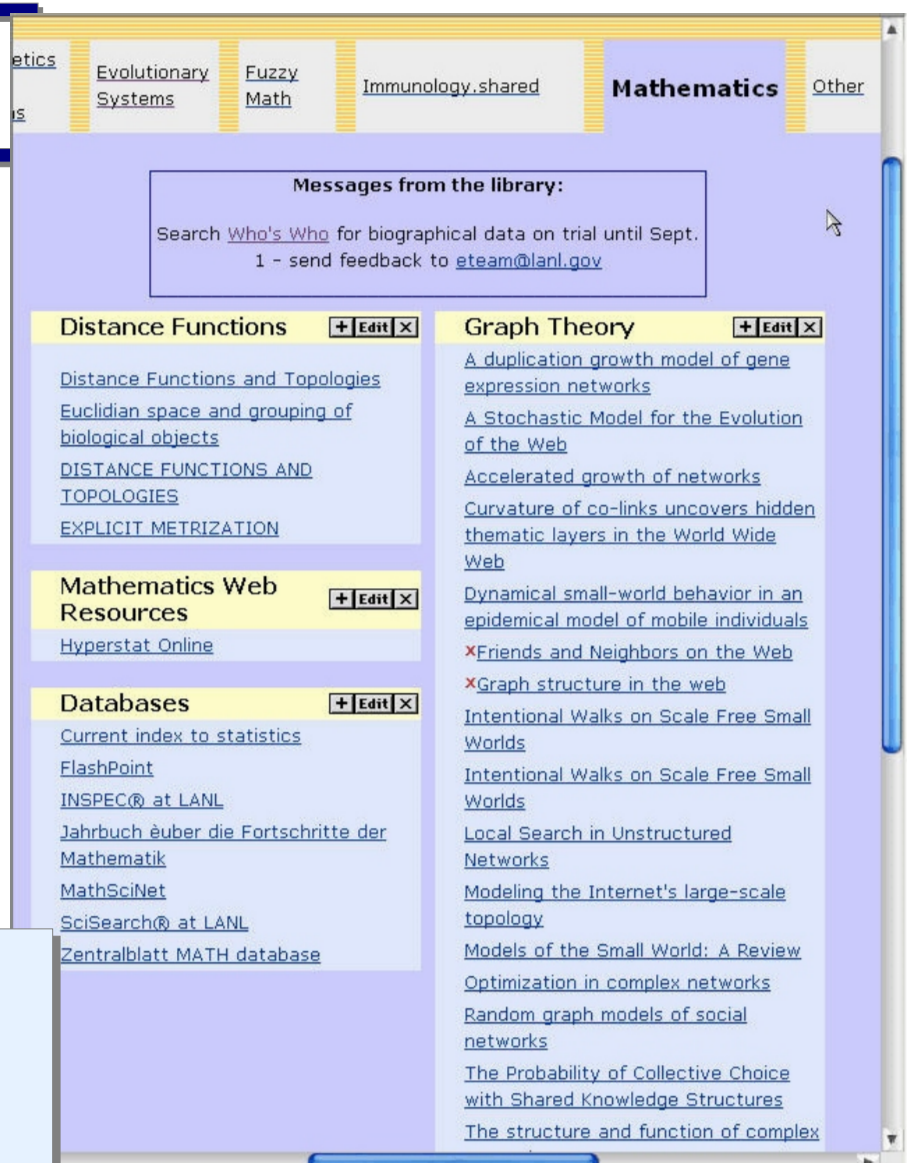
rocha@lanl.gov

MyLibrary

Recommendation Systems



- Three nested entities:
  - Libraries ⇨ Folders ⇨ Links
    - A *library/personality* is associated with a given area of interest and consists of one or more folders.
    - A *folder* contains related types of links within a library
    - A *link* is a URL.



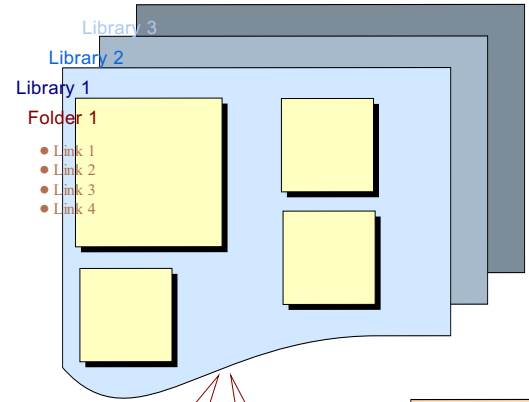
Luis Rocha
2003

rocha@lanl.gov

Journal Extraction

From Links

With Tiago Simas



Personality \ ISSN	
	$A{:}P\times I$

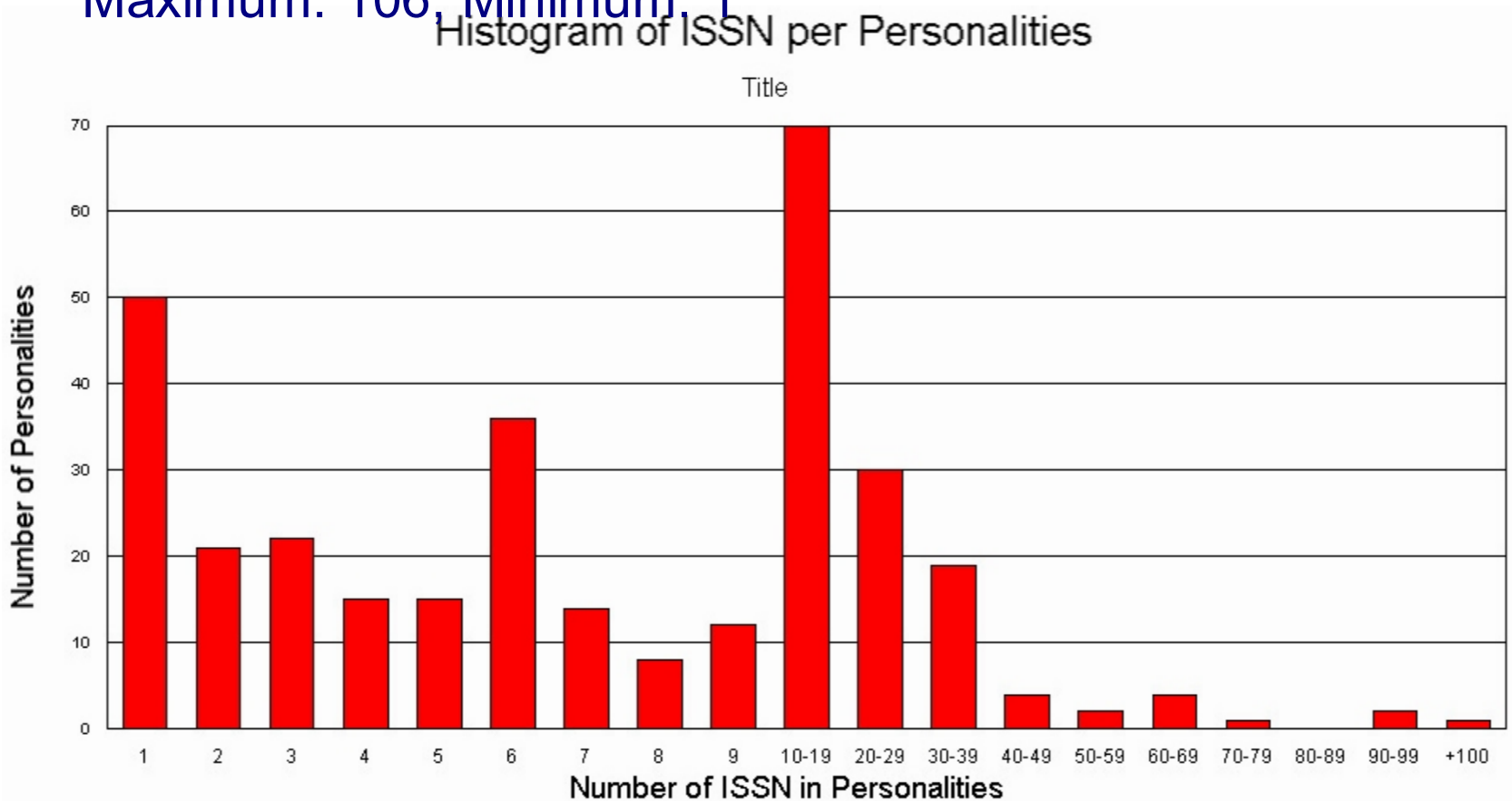
User \ ISSN	
	$B{:}U\times I$

P: set of n_p personalities (326 with at least one ISSN)
U: set of n_u users (253 with at least one ISSN)
I: set of m ISSN (1623 occuring at least once)
$A(p_s, i_t) = a_{s,t}$: number of times ISSN i_t occurs in personality p_s
$N(p_s)$: number of ISSN contained in personality p_s
$B(u_s, i_t) = b_{s,t}$: number of times ISSN i_t occurs in user u_s
$N(u_s)$: number of ISSN contained in user u_s

rocha@lanl.gov

$N(p_s)$: number of ISSN contained in personality p_s

Mean: 12.39, Median: 7, Standard Deviation: 15.06
Maximum: 106, Minimum: 1



15.34% contain *only 1 ISSN*
21.78% contain *no more than 2*
53% contain *no more than 7*
95% contain *less than 30*

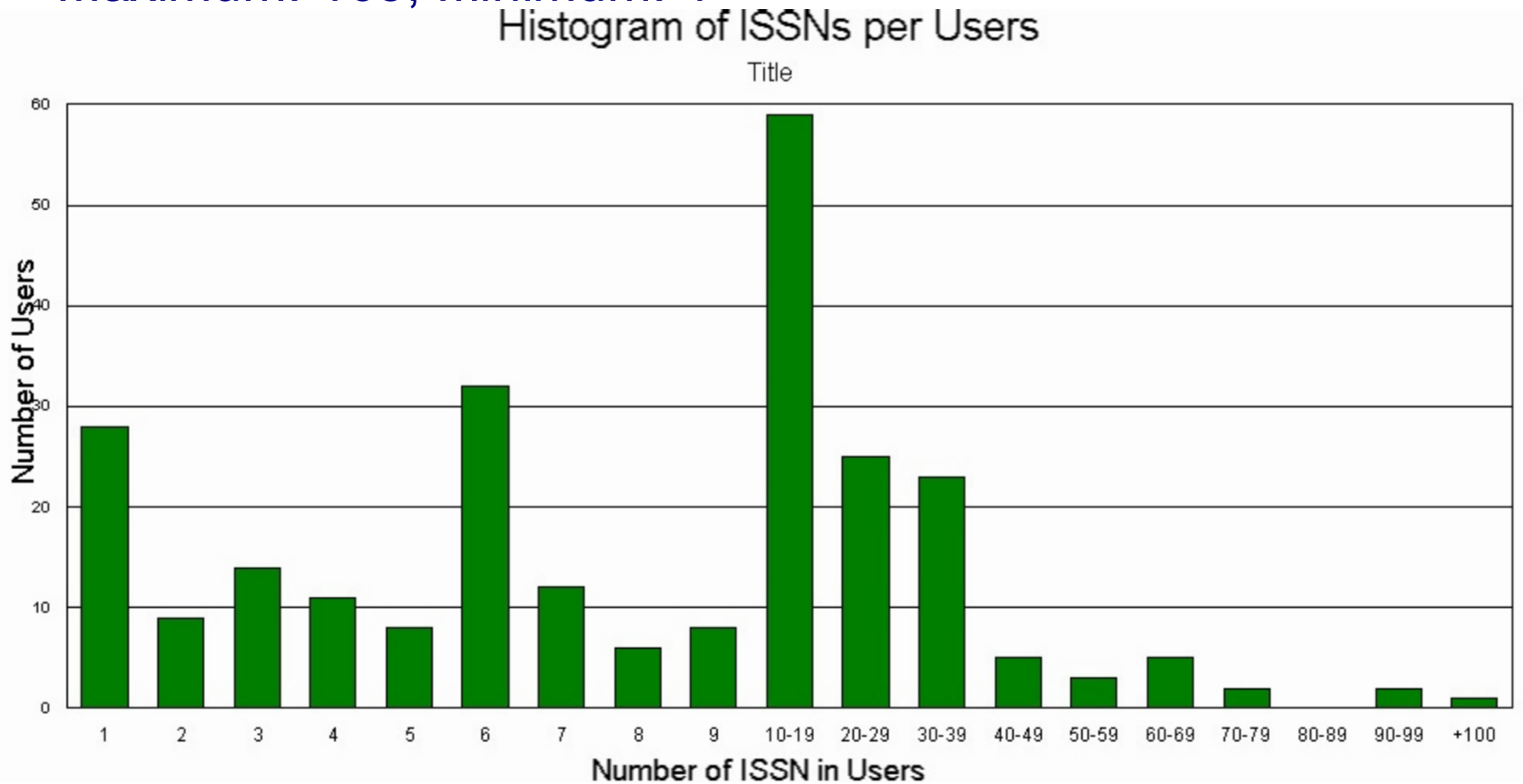
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

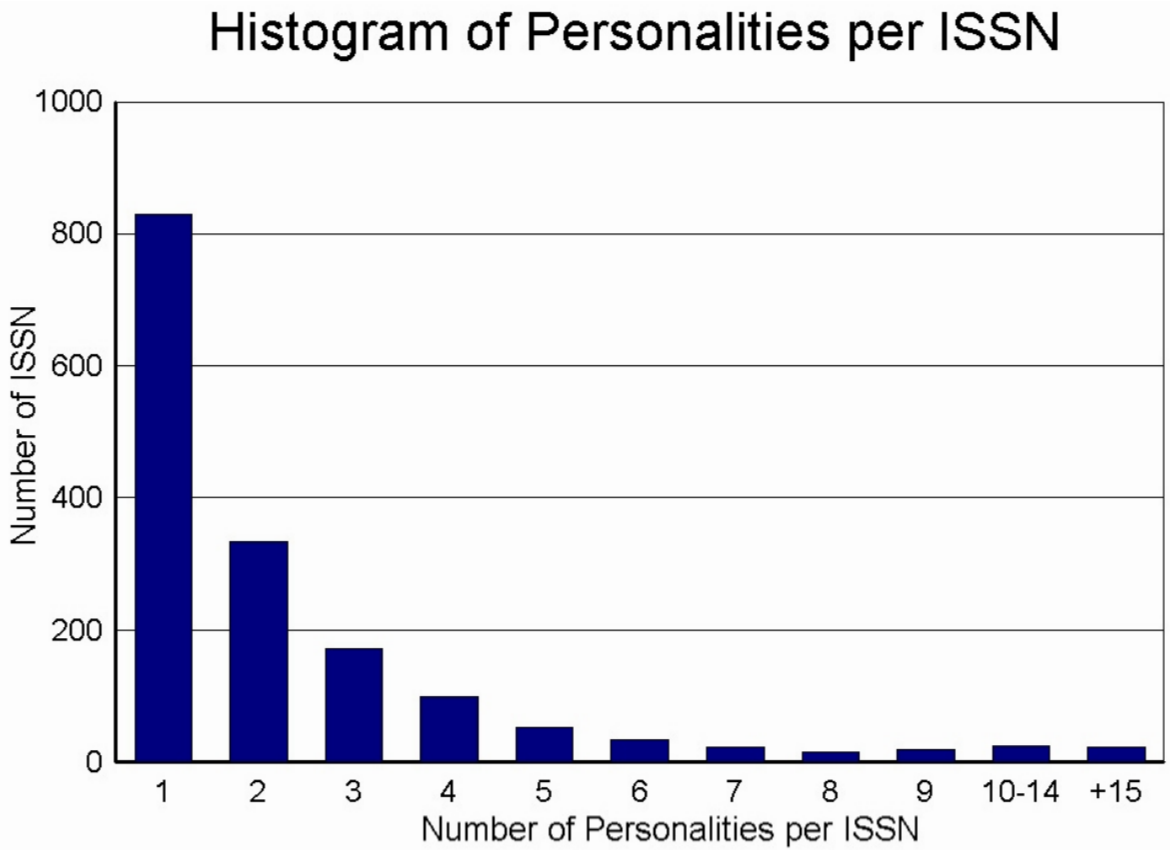
$N(u_s)$: number of ISSN contained in user u_s

Mean: 15.7, Median: 9, Standard Deviation: 18.81
Maximum: 165, Minimum: 1



11.07% of all users contain only *1 ISSN*
14.62% contain *no more than 2*
50.59% contain *no more than 9*
96.05% contain *less than 60*

rocha@lanl.gov



51.08% of ISSN occur in only *one personality*
71.66% occur in *at most two personalities*
95.13% occur in *no more than seven personalities*.

$N(i_t)$: number of personalities with ISSN i_t

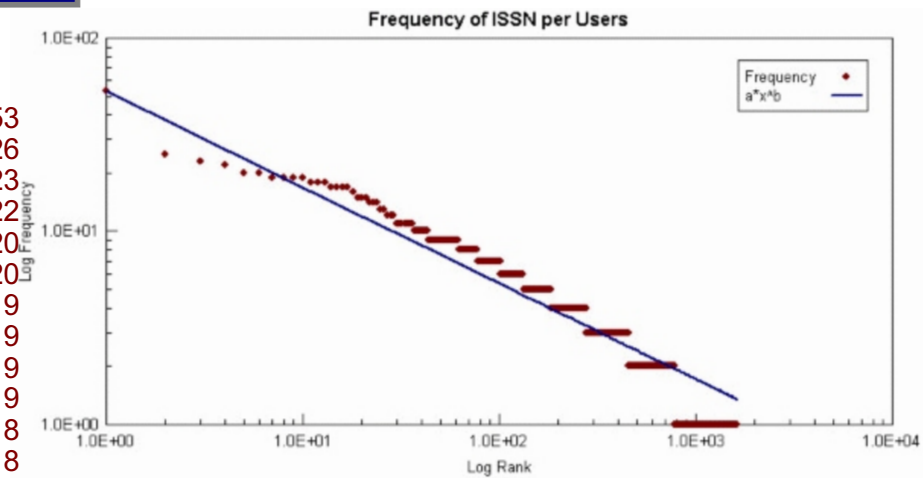
Mean: 2.51, Median: 1, Standard Deviation: 3.04
Maximum: 53, Minimum: 1

rocha@lanl.gov

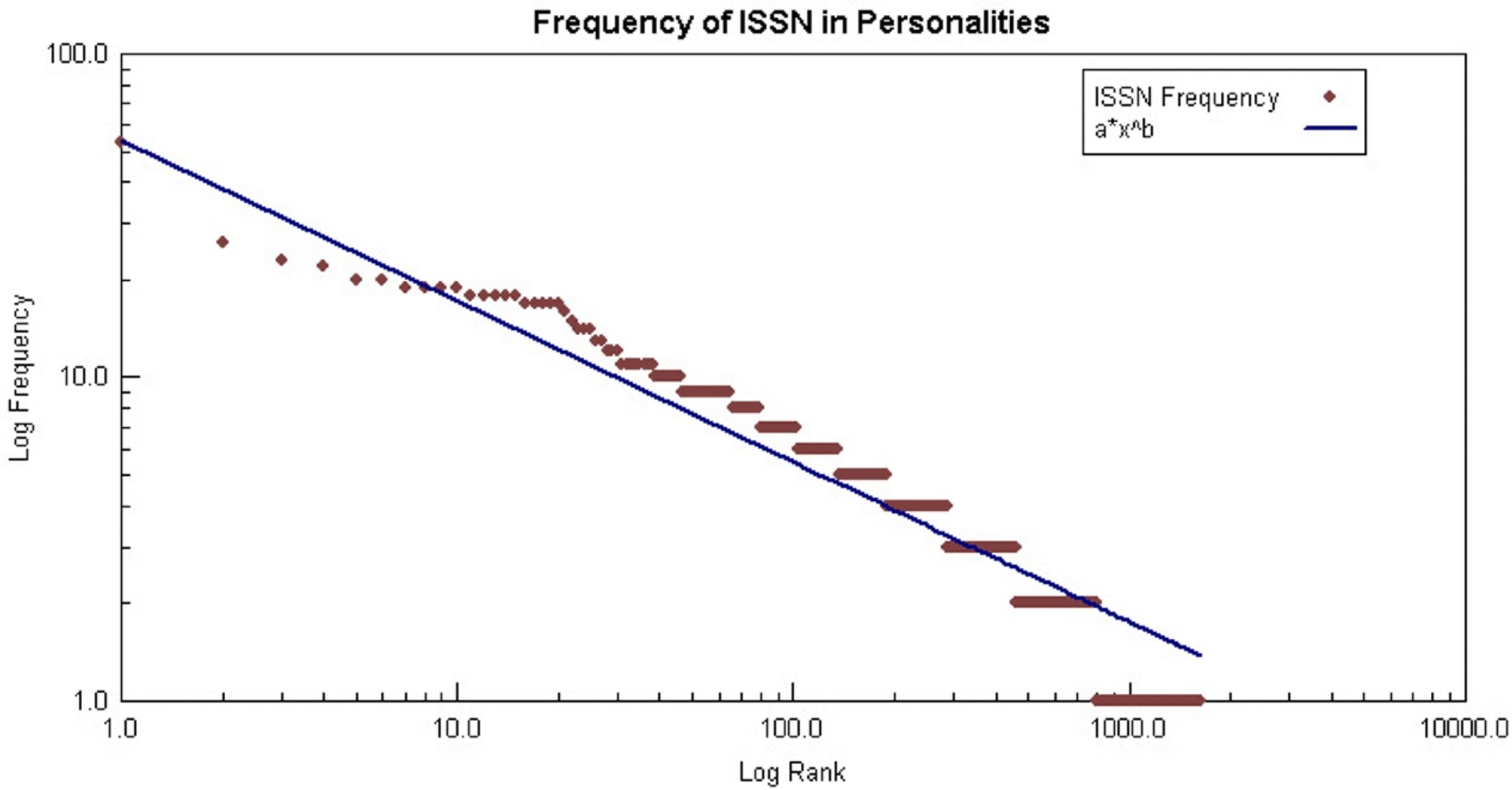
Most Frequent ISSN

Occurences in Personalities

Physical review letters	53
Physical Review B	26
Physical review E	23
Physical review A General physics	22
Journal of physical chemistry B	20
Computers & geosciences	20
Scientific American	19
Journal of the American Chemical Society	19
Journal of Chemical Physics	19
Reviews of modern physics	19
Bioinformatics	18
IEEE trans. on geoscience and remote sensing	18
PNAS	18
Journal of computational physics	18
Advances in water resources	18
Journal of applied geophysics	17
Applied geochemistry	17
APL	17
Journal of physical chemistry A	17
Phil. mag. B Physics of condensed matter	17
Bul.of Environmental Contamination and Toxicology	16
Journal of applied physics	15
American journal of physics	14
Analytical chemistry	14
DLib	14
Chemical physics	13
NIM	13
Chemical physics letters	12
Physics reports	12
Physical review A	12



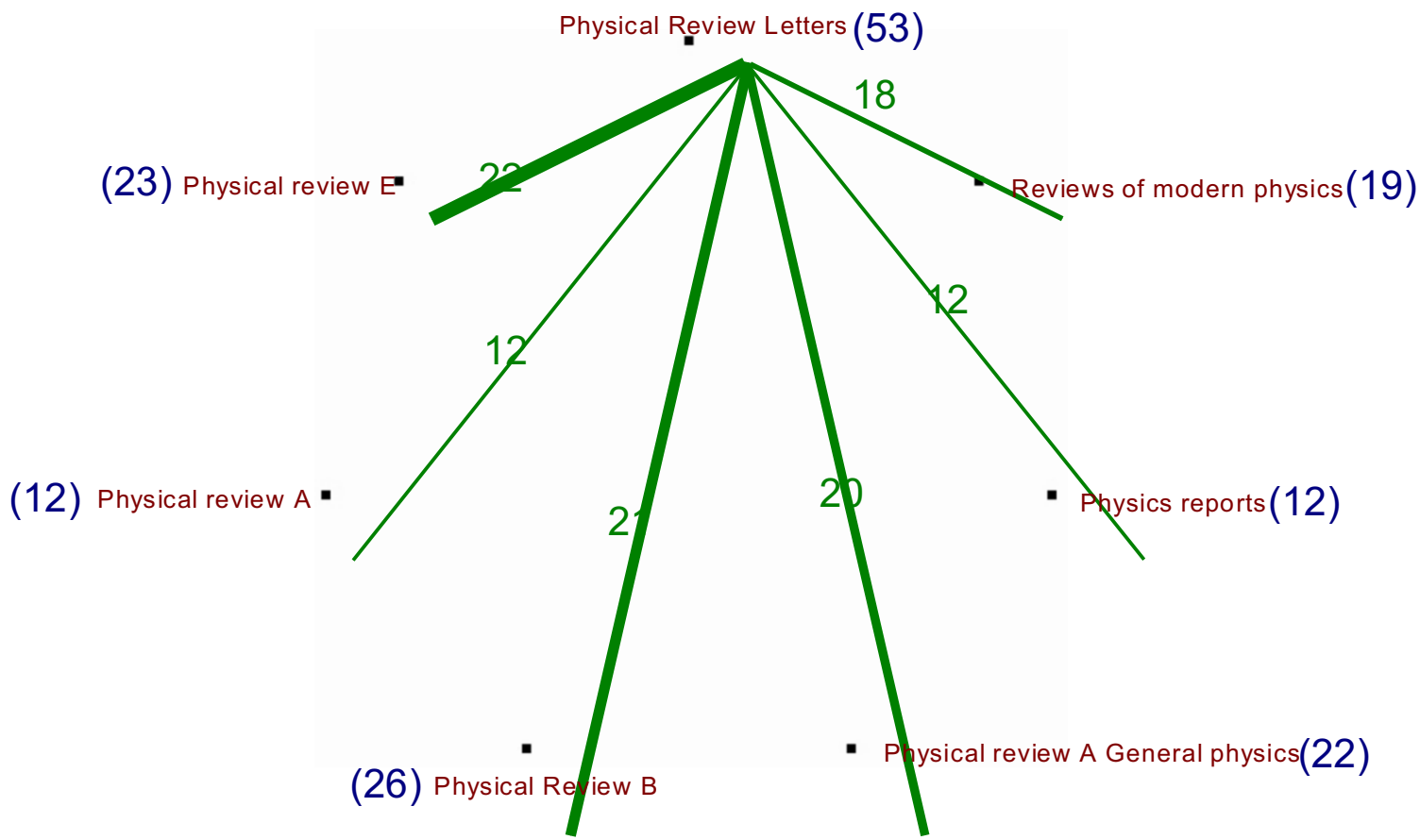
Nature	11
Physics of plasmas	11
Accounts of chemical research	11
Advanced Materials	11
Journal of physics Condensed matter	11
Journal of computational biology	11
Biochemical journal	11
Computer physics communications	11
ACM transactions on modeling and computer simulation	10
Materials science & engineering A	10
Physical review C Nuclear physics	10
Advances in physics	10
Inorganic chemistry	10
Trends in BioTechnology	10
Review of Scientific Instruments	10
Science	10


Frequency of ISSN in Personalities
ISSN Frequency
a*x^b
Log Frequency
100.0
10.0
1.0
1.0
10.0
100.0
1000.0
10000.0
Log Rank

rocha@lanl.gov

How ro Establish Relevance or closeness?

The Neighborhood of a popular Node: Physical Review Letters



Frequency Weights?

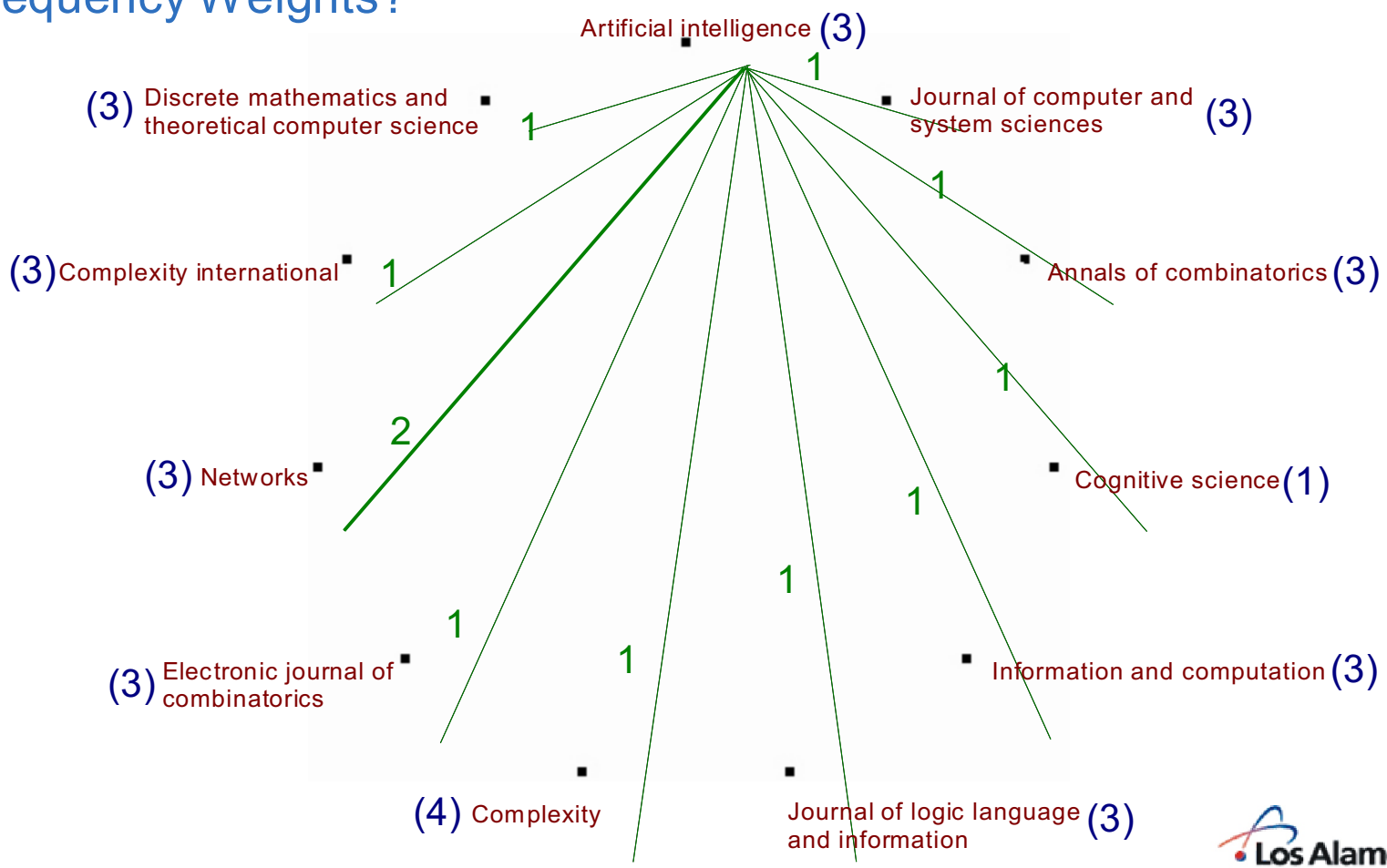
Luis Rocha
2003

rocha@lanl.gov

The Neighborhood of not very Popular Node

Artificial Intelligence

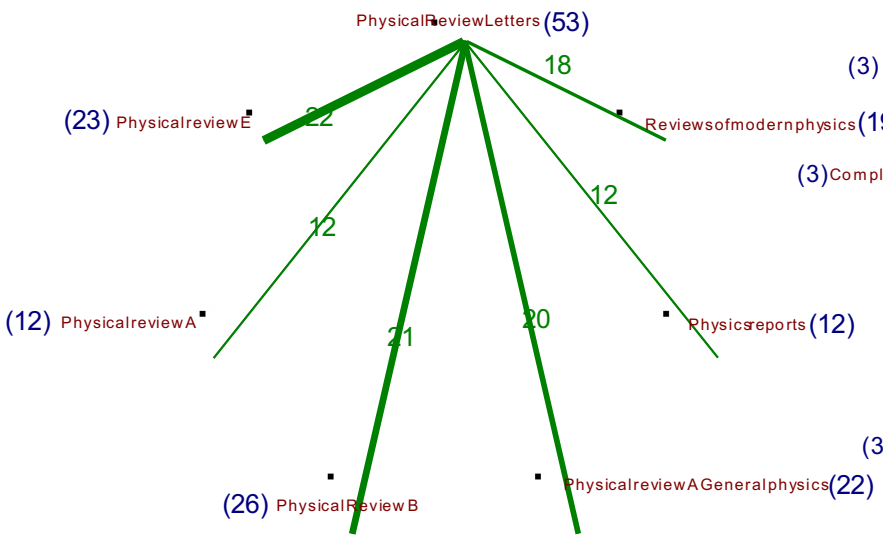
Frequency Weights?

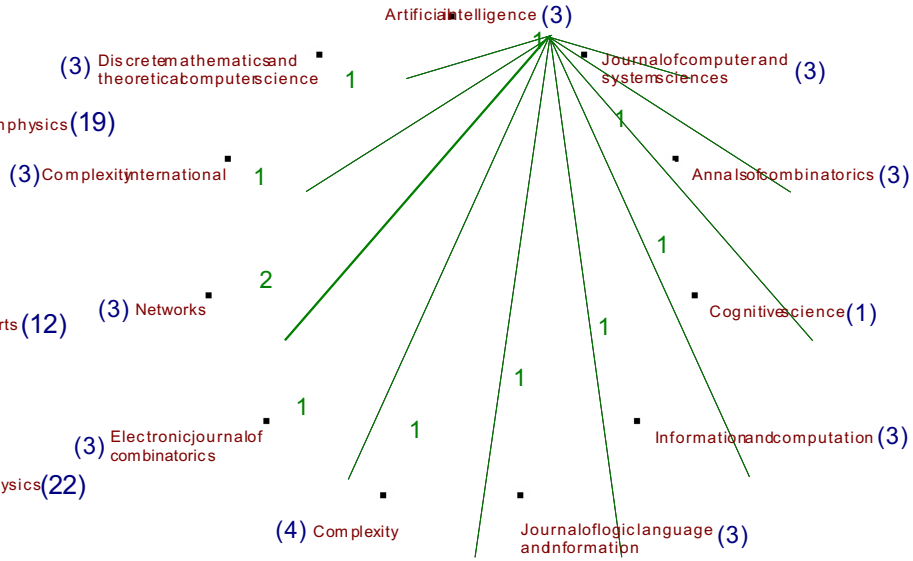


Luis Rocha
2003

rocha@lanl.gov

How to Normalize?





- Normalization
  - By largest Value in Node Neighborhood?
    - Top co-occurence is considered maximum closeness
  - By Total Node Count: Conditional Probability?
    - Not symmetric
  - By co-occurrence probability: Proximity?

rocha@lanl.gov

ISSN and Personality Proximity

	ISSN
Personality	$A{:}P\times I$

Given a binary relation A between sets of Personalities P and ISSN I we extract two proximity relations: $PIP(p_s, p_t)$ is the probability that both personalities p_s and p_t link to the same ISSN $i \in I$. Conversely, $IPP(i_s, i_t)$ is the probability that both ISSN i_s *and* i_t co-occur in the same personality (given that one of them occurs) $p \in P$.

$$pip\ p_s, p_t = \frac{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}}{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}} = \frac{N_{\cap}\ p_s, p_t}{N_{\cup}\ p_s, p_t}$$

(Personality ISSN Proximity)

$$ipp\ (i_s,\ i_t) = \frac{\sum_{k=1}^{m} (a_{i,k} \wedge a_{j,k})}{\sum_{k=1}^{m} (a_{i,k} \vee a_{j,k})} = \frac{N_{\cap}(i_s, i_t)}{N_{\cup}(i_s, i_t)}$$

(ISSN Personality Proximity)

$$XYP(x_i, x_j) = \frac{\sum_{k=1}^{m} (r_{i,k} \wedge r_{j,k})}{\sum_{k=1}^{m} (r_{i,k} \vee r_{j,k})}; \quad YXP(y_i, y_j) = \frac{\sum_{k=1}^{n} (r_{k,i} \wedge r_{k,j})}{\sum_{k=1}^{n} (r_{k,i} \vee r_{k,j})}$$

Measuring "Knowledge" Associations

From Document Relations

rocha@lanl.gov

- Document × Document
  - Co-Citation or Hyperlink structure
- Document × Keyterms
  - Keyterm Co-Occurrence
- Document/Dataset × Gene Expression
  - Gene Co-Occurrence or Co-Expression
- Document × Author
  - Co-Authorship (Collaboration Network)

Given a binary relation R between sets X and Y we extract two proximity relations: $XYP(x_i, x_j)$ is the probability that both x_i and x_j are related in R to the same element $y \in Y$. Conversely, $YXP(y_i, y_j)$ is the probability that both y_i *and* y_j are related in R to the same element $x \in X$.

rocha@lanl.gov

Pointwise Mutual Information

Other non-proximity Probability Measures of Association

Conditional Probability

Given a binary relation R between sets X and Y we extract :

$$P_X(x_i, x_j) = P(x_i | x_j) = \frac{\sum_{k=1}^{m} (r_{i,k} \wedge r_{j,k})}{\sum_{k=1}^{m} (r_{j,k})}; \quad P_Y(y_i, y_j) = P(y_i | y_j) = \frac{\sum_{k=1}^{n} (r_{k,i} \wedge r_{k,j})}{\sum_{k=1}^{n} (r_{k,j})}$$

$P_X(x_i, x_j)$ is the probability that x_i is related in R to $y \in Y$, given that x_j is related to y.

$P_Y(y_i, y_j)$ is the probability that y_i is related in R to $x \in X$, given that x_y is related to x.

rocha@lanl.gov

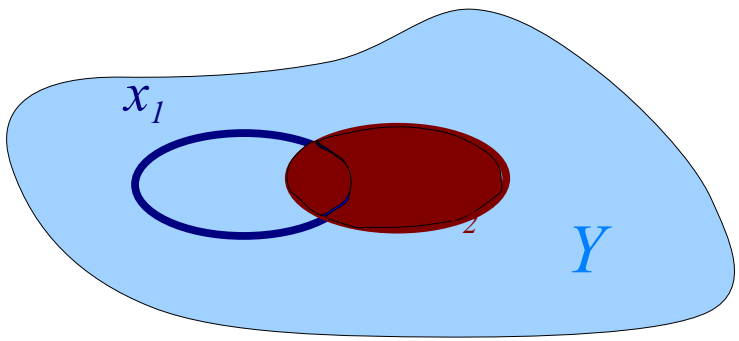
Proximity vs. Conditional Probability

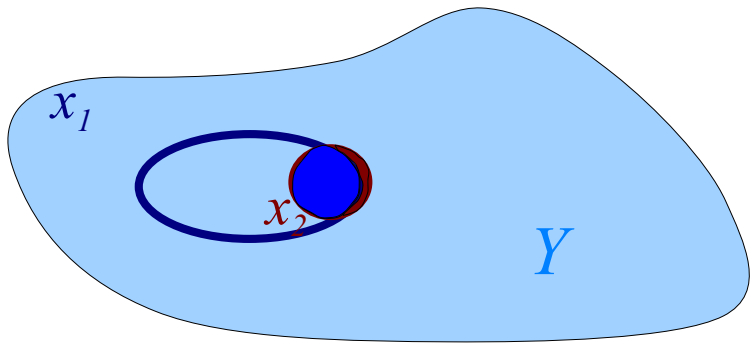
- P_X and P_Y are not symmetric
  - can measure a strong degree of association between two elements, when that association is one-sided only.
  - In many applications, when we think of a strong association between two elements, we expect both directions of association to be similar.
  - Distances are symmetric, proximity is the semantic inverse of distance.

rocha@lanl.gov

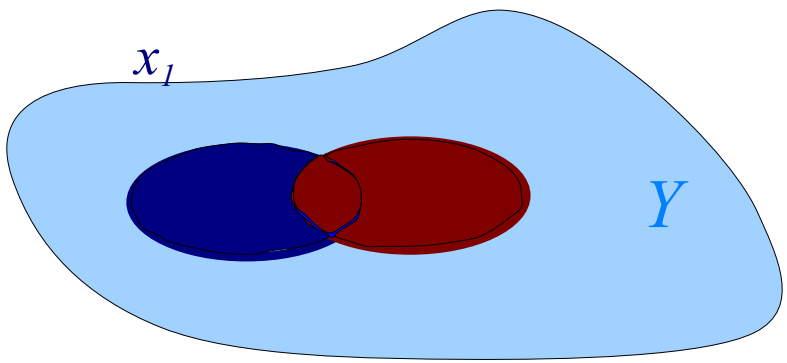
Conditional Probability and Proximity Measures

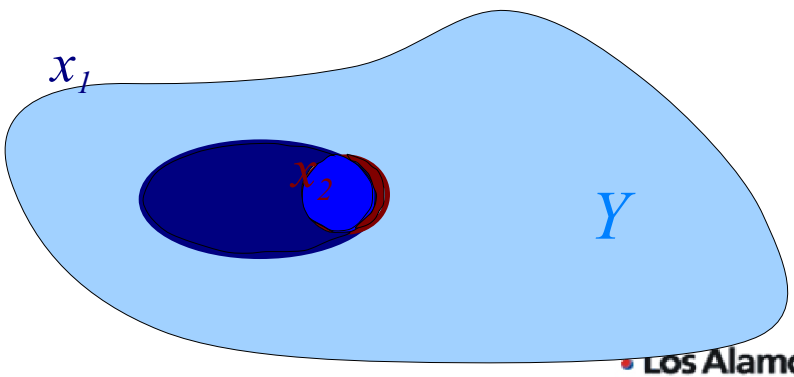
Conditional Probability





Proximity





rocha@lanl.gov

Using Information Retrieval

Immunology Testcase

With Andreas Rechtsteiner

Cytokines

IL-2 OR interleukin-2
IL-3 OR interleukin-3
IL-4 OR interleukin-4
IL-6 OR interleukin-6
IL- 7 OR interleukin-7
IL-8 OR interleukin-8
IL-10 OR interleukin-10
IL-12 OR interleukin- 12
IL-13 OR interleukin-13
IL-15 OR interleukin-15
GM-CSF
IFNgamma
TNFalpha
MCP-1

Receptor Molecules

CD25 OR tac
CD122
CD132 OR "common gamma chain"
CD123
beta
CD124
CD126
CD130 OR gp130
CD127
CXCR1 OR Cdw128a
CXCR2 OR Cdw128b
Cdw210
CD212
CD213a1
CD213a2
CD116
CD119
CD120a
CD120b
CCR2b

Signaling molecules: their levels affect response

rocha@lanl.gov

Pointwise Mutual Information

Information Retrieval

Turney, P.D. (2001). "Mining the Web for Synonyms: PMI-IR versus LSA on TOEFL." *Proceedings of the Twelfth European Conference on Machine Learning* (ECML-2001), Freiburg, Germany, pp. 491-502. http://extractor.iit.nrc.ca/reports/ECML2001.html

- PMI-IR developed to deal with Synonyms
  - Problem word and set of choice words for synonyms (on TOEFL and ESL)
    - Problem: ***levied***
    - Choices: *imposed*, *believed*, *requested*,*correlated*
  - Assigns a *co-occurrence* score to each choice, and selects the choice that maximizes the score.
    - $\text{score}(\text{choice}_i) = p(\text{problem} \mid \text{choice}_i) = \text{p}(\text{problem AND choice}_i) / p(\text{choice}_i)$
    - In PMI-IR, the probabilities are calculated using IR from such sources as Altavista.

rocha@lanl.gov

PMI-IR

1. Conditional Probability with Altavista AND (when both terms appear in the same document)

4 Scores using Altavista

$$score_4(choice_i) = \frac{hits((problem \text{ NEAR } choice_i) \text{ AND } context \text{ AND NOT}((problem \text{ OR } choice_i) \text{ NEAR "not"}))}{hits(choice_i \text{ AND } context \text{ AND NOT}(choice_i \text{ NEAR "not"}))}$$

2. Conditional probability with Altavista NEAR (when both terms appear within 10 words of each other in the same document)

3. Tends to reduce equal scores for synonyms and antonyms

$$score_3(choice_i) = \frac{hits(problem \text{ NEAR } choice_i) \text{ AND NOT}((problem \text{ OR } choice_i) \text{ NEAR "not"})}{hits(choice_i \text{ AND NOT}(choice_i \text{ NEAR "not"}))}$$

4. Accounts for context words

$$score_2(choice_i) = \frac{hits(problem \text{ NEAR } choice_i)}{hits(choice_i)}$$

LSA Problems: "Every year in the early spring farmers [tap] maple syrup from their trees (drain; boil; knock; rap)." The problem word *tap*, out of context, might seem to best match the choice words *knock* or *rap*, but the context maple syrup makes *drain* a better match for tap.

$$score_1(choice_i) = \frac{hits(problem \text{ AND } choice_i)}{hits(choice_i)}$$

$$context = \{context_1, context_2, \ldots, context_m\}$$

We use NEAR

Luis Rocha
2003

rocha@lanl.gov

Proximity Queries

$$prox_3(problem, choice_i) = \frac{hits((problem \text{ NEAR } choice_i) \text{ NEAR } context)}{hits(problem \text{ OR } choice_i)}$$

$$prox_2(problem, choice_i) = \frac{hits(problem \text{ NEAR } choice_i)}{hits(problem \text{ OR } choice_i)}$$

$$prox_1(problem, choice_i) = \frac{hits(problem \text{ AND } choice_i)}{hits(problem \text{ OR } choice_i)}$$

NEAR: co-occurrence within 10 words

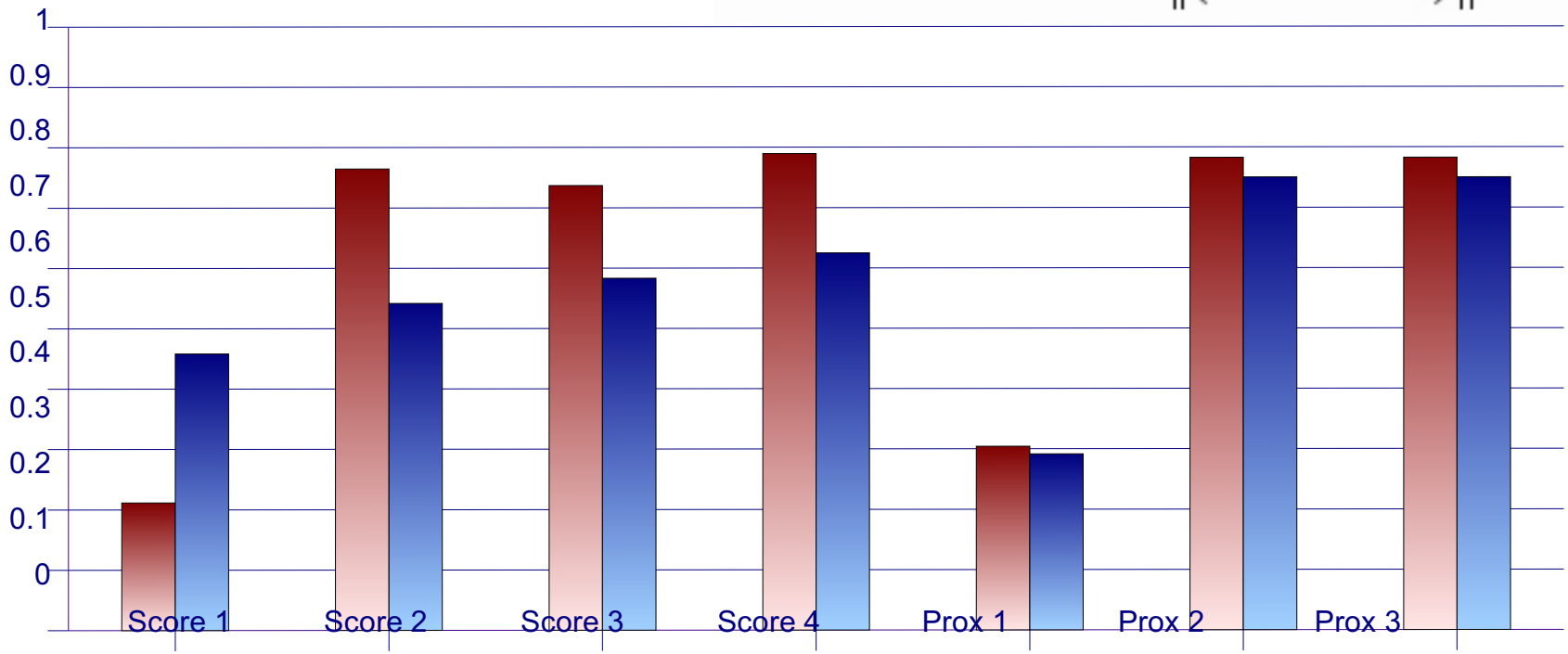
context = {receptor}

rocha@lanl.gov

Precision and Recall

$$precision = \frac{\|\{relevant\} \cap \{retrieved\}\|}{\|\{retrieved\}\|}$$

Values Obtained with FaCSO



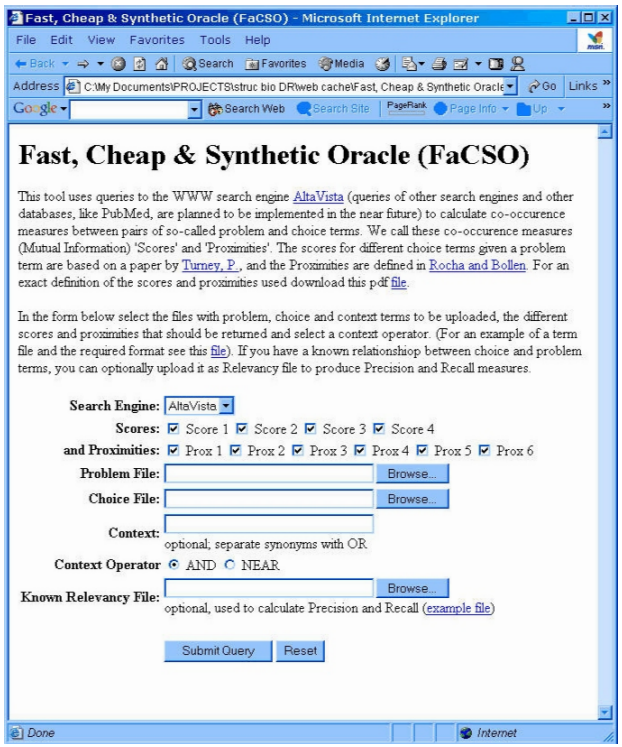
Precision: probability that an identified association is relevant

Recall: probability that an association has been identified given that it is relevant

rocha@lanl.gov

Measuring Associations

Using Altavista



Rocha, Luis M. and Andreas Rechtsteiner [2003]. "Mining the Web with Proximity Measures". In Preparation.

- Discovers Relevant Associations
- Retrieves Documents substantiating the associations
- Being Developed for PubMed

Fast, Cheap & Synthetic Oracle (FaCSO)

http://bimmer.c3.lanl.gov/~andreas/easyfast8.html

Cytokine Set

Los Alamos NATIONAL LABORATORY

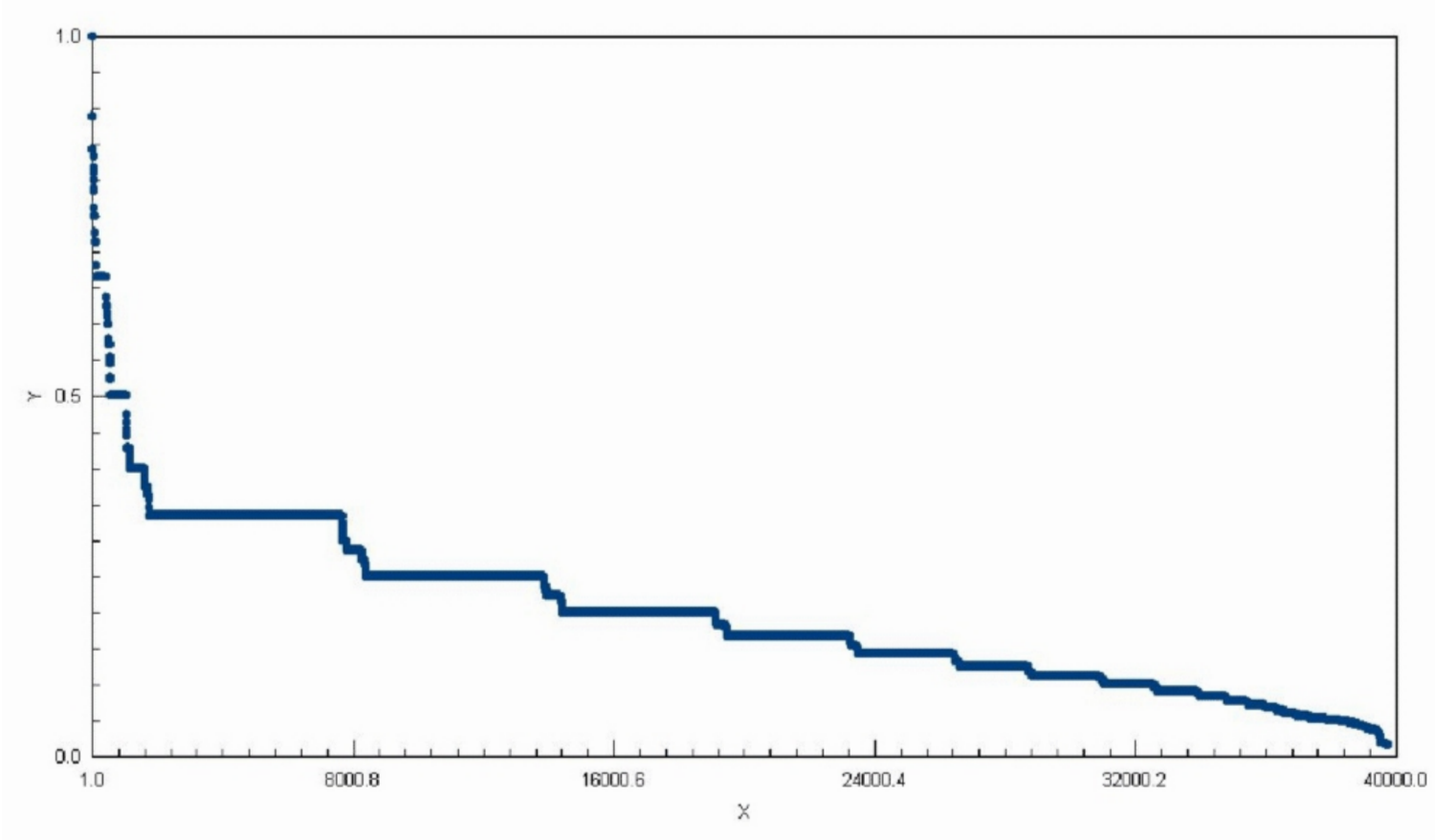
rocha@lanl.gov

ISSN Personality Proximity

$$ipp\ i_s, i_t \quad \frac{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}\ N_{\cap}\ i_s, i_t}{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}\ N_{\cup}\ i_s, i_t}$$

(ISSN Personality Proximity)

IPP: exponential distribution of weights



Luis Rocha
2003

rocha@lanl.gov

IPP

m=1623

$$ipp\ i_s, i_t \quad \frac{N_\cap\ i_s, i_t}{N_\cup\ i_s, i_t} \geq 3$$

$ipp \geq 0.3$

0031-9007--Physical review letters
- 1095-3787 --Physical review E: 0.4074
- 0556-2805--Physical Review B: 0.3621
- 0034-6861--Reviews of modern physics: 0.3333
- 0556-2791--Physical review A General physics: 0.3636

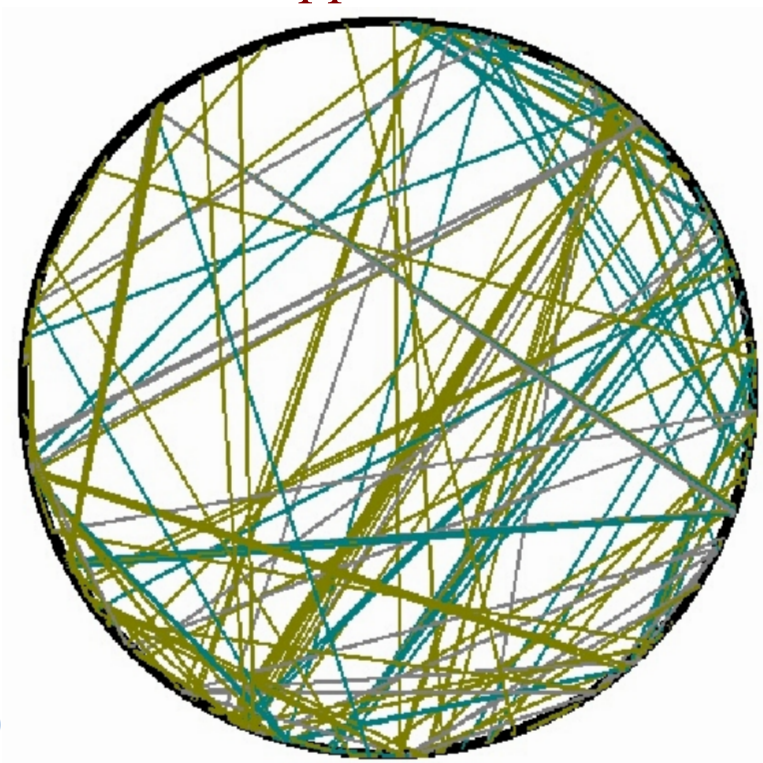
0556-2805--Physical Review B
- 0031-9007--Physical review letters: 0.3621
- 0370-1573--Physics reports: 0.3103
- 0921-4526--Physica B Condensed matter: 0.3462
- 0921-4534--Physica C Superconductivity: 0.3462
- 0034-6861--Reviews of modern physics: 0.3235
- 0556-2791--Physical review A General physics: 0.3333
- 1434-6036--European physical journal B: 0.3077

1095-3787--Physical review E
- 0031-9007--Physical review letters: 0.4074

0556-2791--Physical review A General physics
- 0031-9007--Physical review letters: 0.3636
- 0370-1573--Physics reports: 0.3077
- 0556-2805--Physical Review B: 0.3333
- 1089-490x --Physical review C Nuclear physics: 0.3913
- 0034-6861--Reviews of modern physics: 0.4643
- 1434-6036--European physical journal B: 0.3043

1089-5647--Journal of physical chemistry B
- 0002-7863--Journal of the American Chemical Society: 0.3000
- 0021-9606--Journal of Chemical Physics: 0.6250
- 1089-5639--Journal of physical chemistry A: 0.7619
- 0009-2614--Chemical physics letters: 0.6000
- 0301-0104--Chemical physics: 0.5714
- 0743-7463--Langmuir: 0.3810

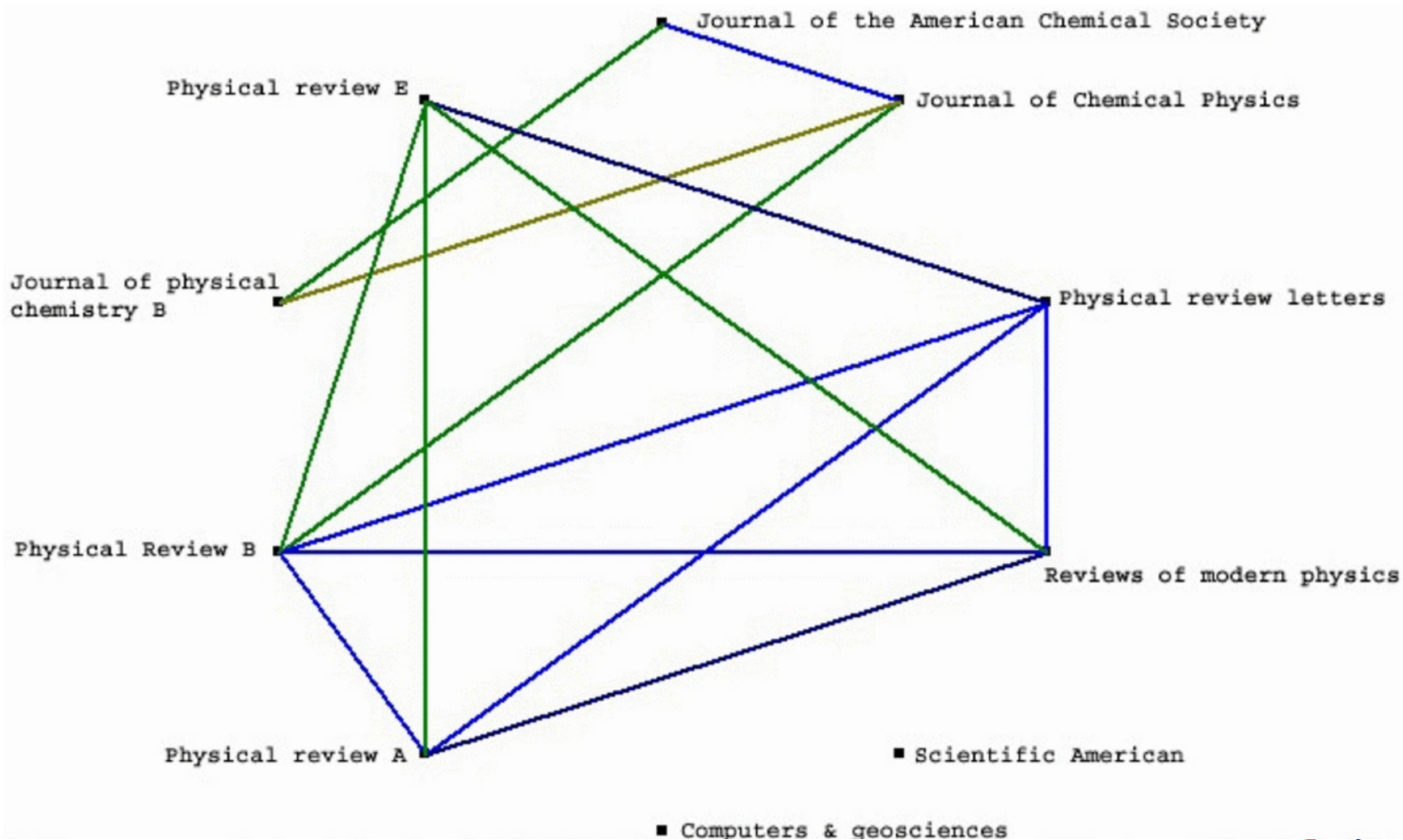
$ipp \geq 0.5$



rocha@lanl.gov

IPP SubGraph for 10 most Frequent ISSN

ipp ≥ 0.2

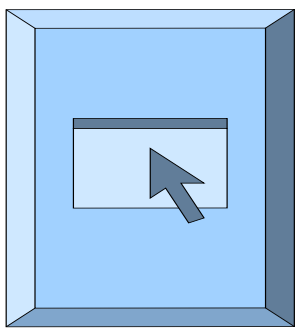


rocha@lanl.gov

Recommendations based on Proximity

For MyLibrary

- IPP
  - Recommendations of ISSN based on co-occurrence in Personalities
    - Users who linked to this journal, also linked to...
- PIP
  - Recommendations of other users' personalities: collaboration
    - These personalities are similar to yours
  - Recommendations of specific links in close personalities
    - Users who read many of the same journals where interested in these links



Rocha, Luis M. and Johan Bollen [2001]. "Biologically Motivated Distributed Designs for Adaptive Knowledge Management". In: *Design Principles for the Immune System and other Distributed Autonomous Systems*. L. Segel and I. Cohen (Eds.) Santa Fe Institute Series in the Sciences of Complexity. Oxford University Press, pp. 305-334.

Rocha, Luis M. [2002]. "Combination of Evidence in Recommendation Systems Characterized by Distance Functions". In: *Proceedings of the 2002 World Congress on Computational Intelligence: FUZZ-IEEE'02*. Honolulu, Hawaii, May 2002. IEEE Press, pp. 203-208.

Rocha, Luis M. and Tiago Simas [2003]. "Recommendation Based on Proximity Measures for a Library Portal". In Preparation.

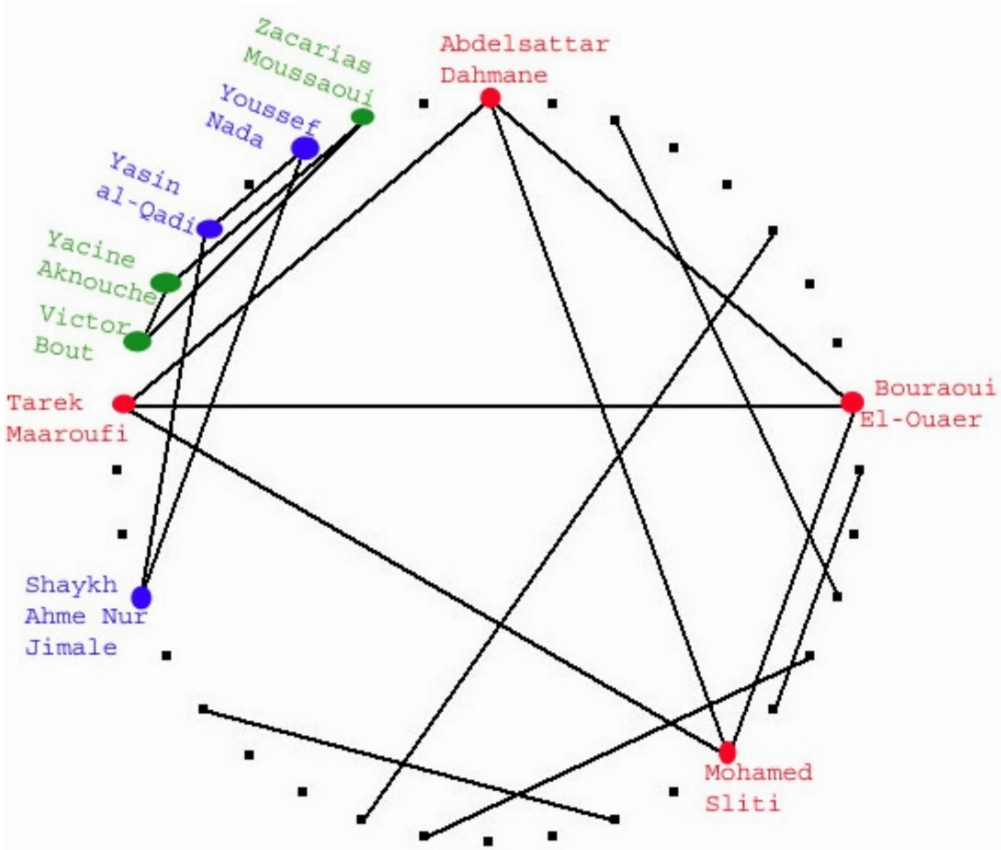
People Network

Rocha, Luis M. [2002]. "Proximity and Semi-Metric Analysis of Social Networks". LAUR 02-6557

rocha@lanl.gov

Detail

Voss, Susan and Cliff Joslyn (2002): ``Advanced Knowledge Integration in Assessing Terrorist Threats', LAUR 02-7867
ftp://ftp.c3.lanl.gov/pub/users/joslyn/knowint.pdf

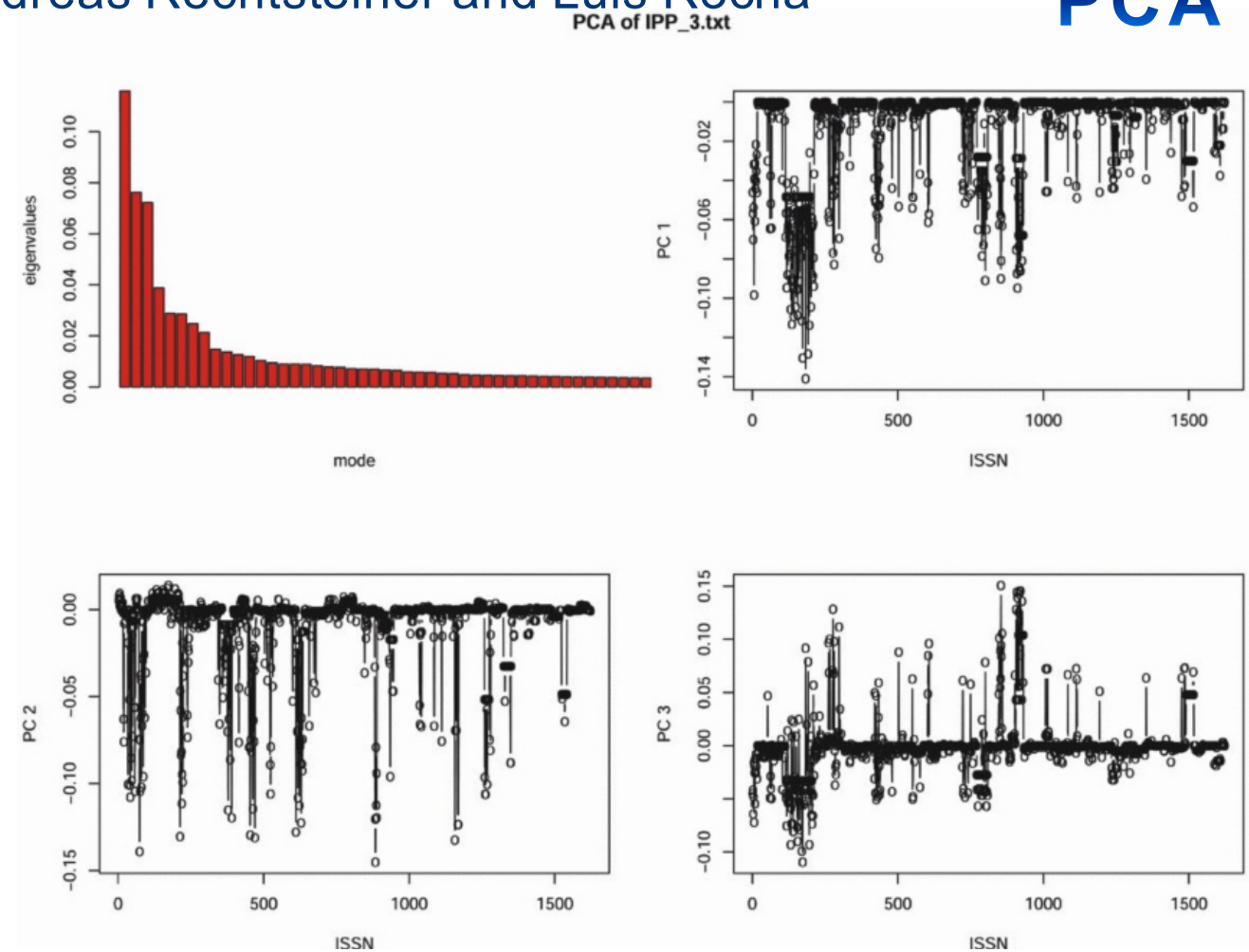


Eigen Analysis of MyLibrary

Andreas Rechtsteiner and Luis Rocha

rocha@lanl.gov

PCA of IPP_3

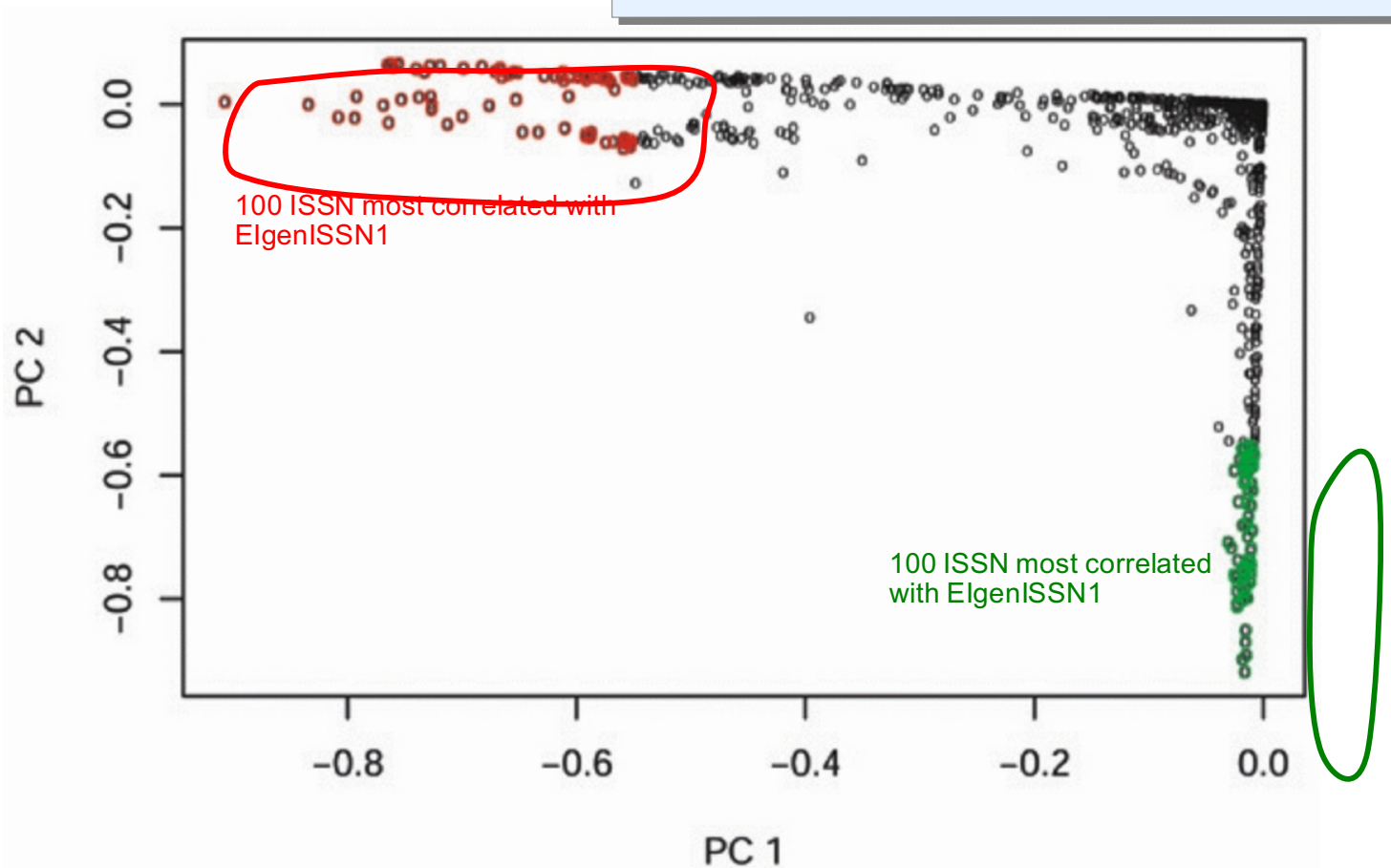


rocha@lanl.gov

PCA of IPP_3

First 2 EigenISSN

- ISSN correlated with either one of the first two EigenISSN, or with neither
  - Very few ISSN correlated with both.
  - 2 very clear main groups



rocha@lanl.gov

Top 30 ISSN most Correlated with EigenISSN 1 and 2

EigenISSN 1: Chemistry and Materials Science

Nanostructured materials 142
Journal of materials chemistry 160
International Journal of Solids and Structures 192
International journal of refractory metals & hard materials 183
Advanced functional materials 136
Applied organometallic chemistry 211
Journal of organic chemistry 154
Annual review of materials science 210
Polyhedron 156
Nano Letters 118
Tetrahedron 196
Topics in current chemistry 173
Radiation effects and defects in solids 121
European journal of inorganic chemistry 6
International journal of inorganic materials 150
Progress in solid state chemistry 202
Chemical reviews 133
Biophysical chemistry 140
Advanced Materials 198
Nuclear energy 187
Advanced engineering materials 135
Journal of Raman spectroscopy 162
Intermetallics 800
Chemistry of materials 425
Chemical Society reviews 134
Journal of biomolecular structure & dynamics 137
Bioorganic chemistry 139
Journal of biochemical and biophysical methods 206
Journal of inorganic biochemistry 170
Chemphyschem 789

EigenISSN 2: Computer Science and Applied Mathematics

Algorithmica 74
Science of computer programming 470
Discrete applied mathematics 884
ACM journal of experimental algorithmics 213
Mathematical programming 629
Applied mathematics and computation 80
ACM transactions on modeling and computer simulation 87
Computer languages 887
Journal of the Association for Computing Machinery 454
SIAM journal on computing 222
Mathematical and computer modelling 618
Mathematics of computation 611
ACM computing surveys 58
ACM letters on programming languages and systems 219
Computers & mathematics with applications 623
Journal of computational and applied mathematics 378
Theoretical computer science 889
IEEE computational science & engineering 76
Computational mathematics and modeling 524
SIAM journal on scientific computing 220
Applied scientific research 1261
ACM trans. on programming languages and systems 1263
Journal of graph algorithms and applications 883
Discrete mathematics 885
IEEE computer graphics and applications 1167
Computers & graphics 1168
Computer graphics 1169
ACM transactions on graphics 47
Communications on pure and applied mathematics 391
Computers & security 41

Los Alamos NATIONAL LABORATORY

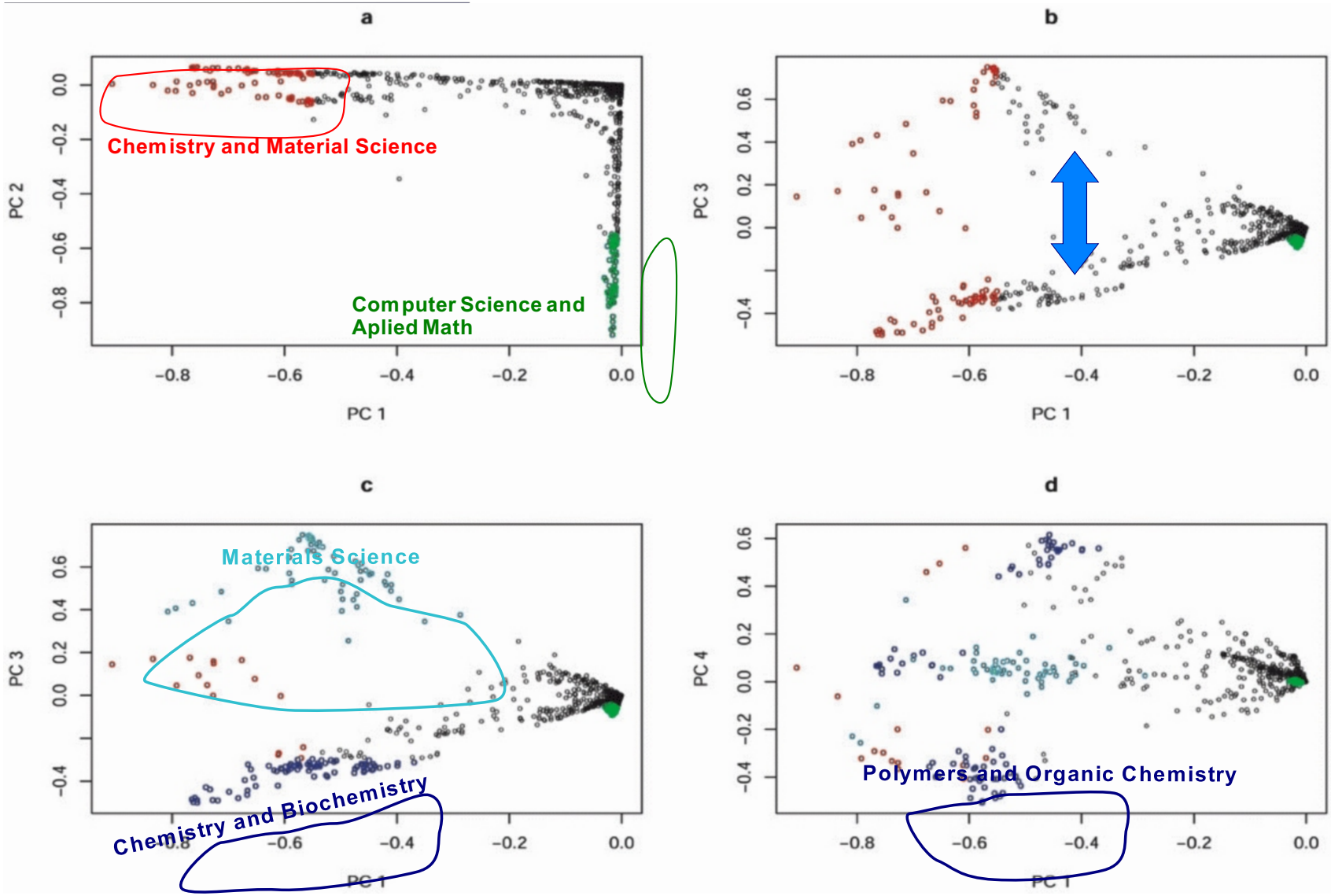
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

PCA of IPP_3

EigenISSN1 with other EigenISSN

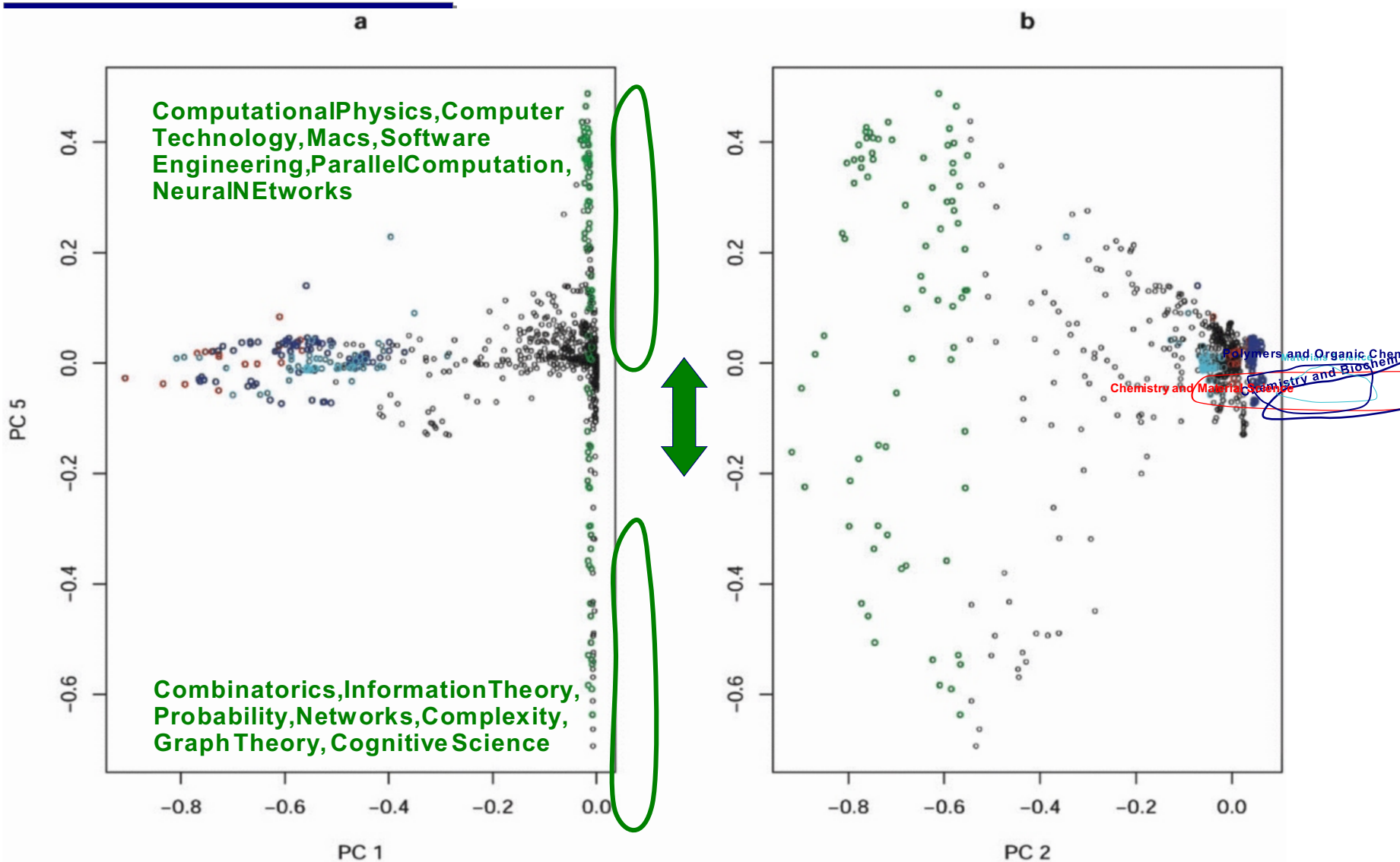


Luis Rocha
2003

PCA of IPP_3

EigenISSN2

rocha@lanl.gov



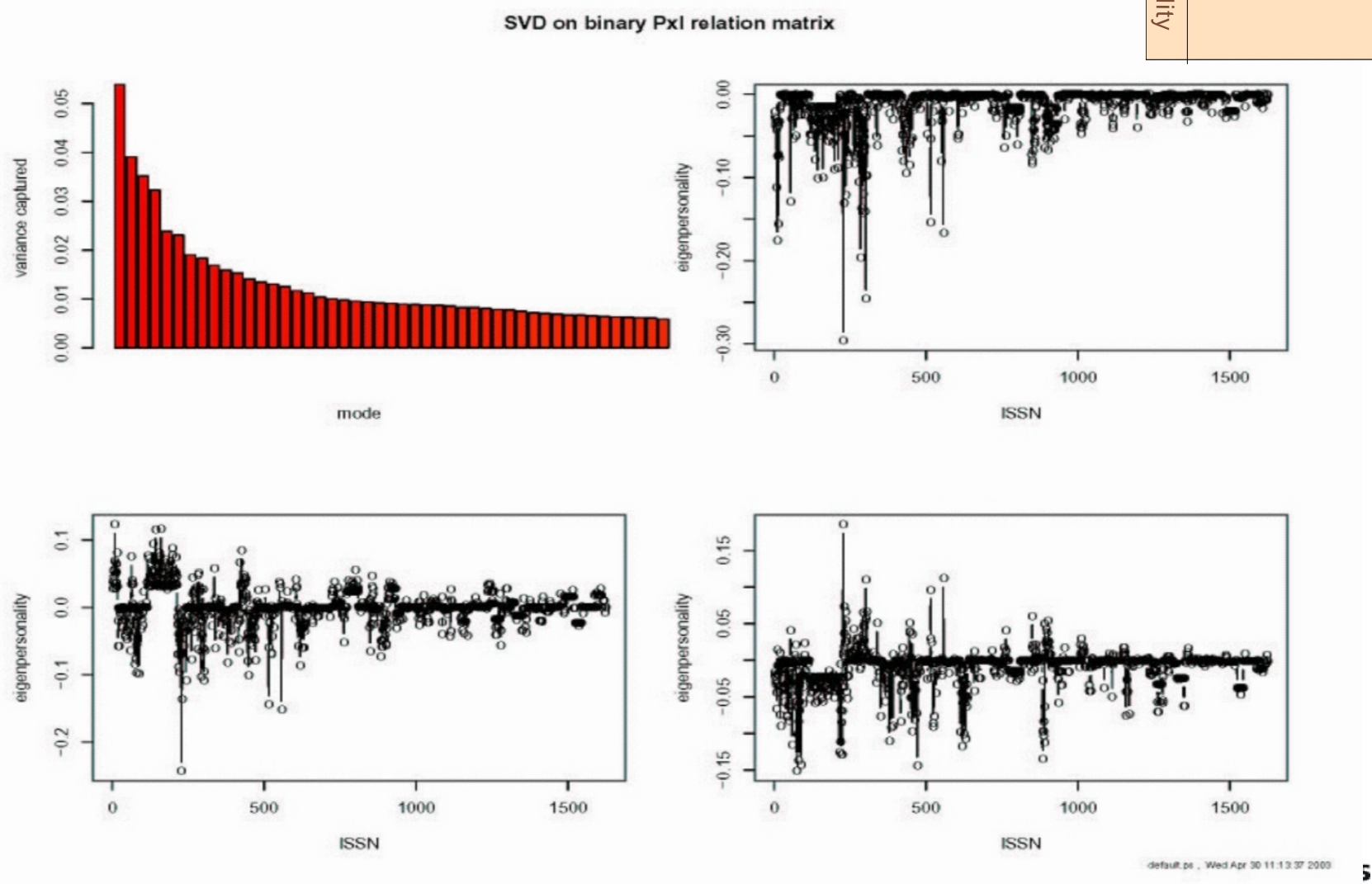
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

SVD of Personality ISSN Relation

	ISSN
Personality	$A: P \times I$

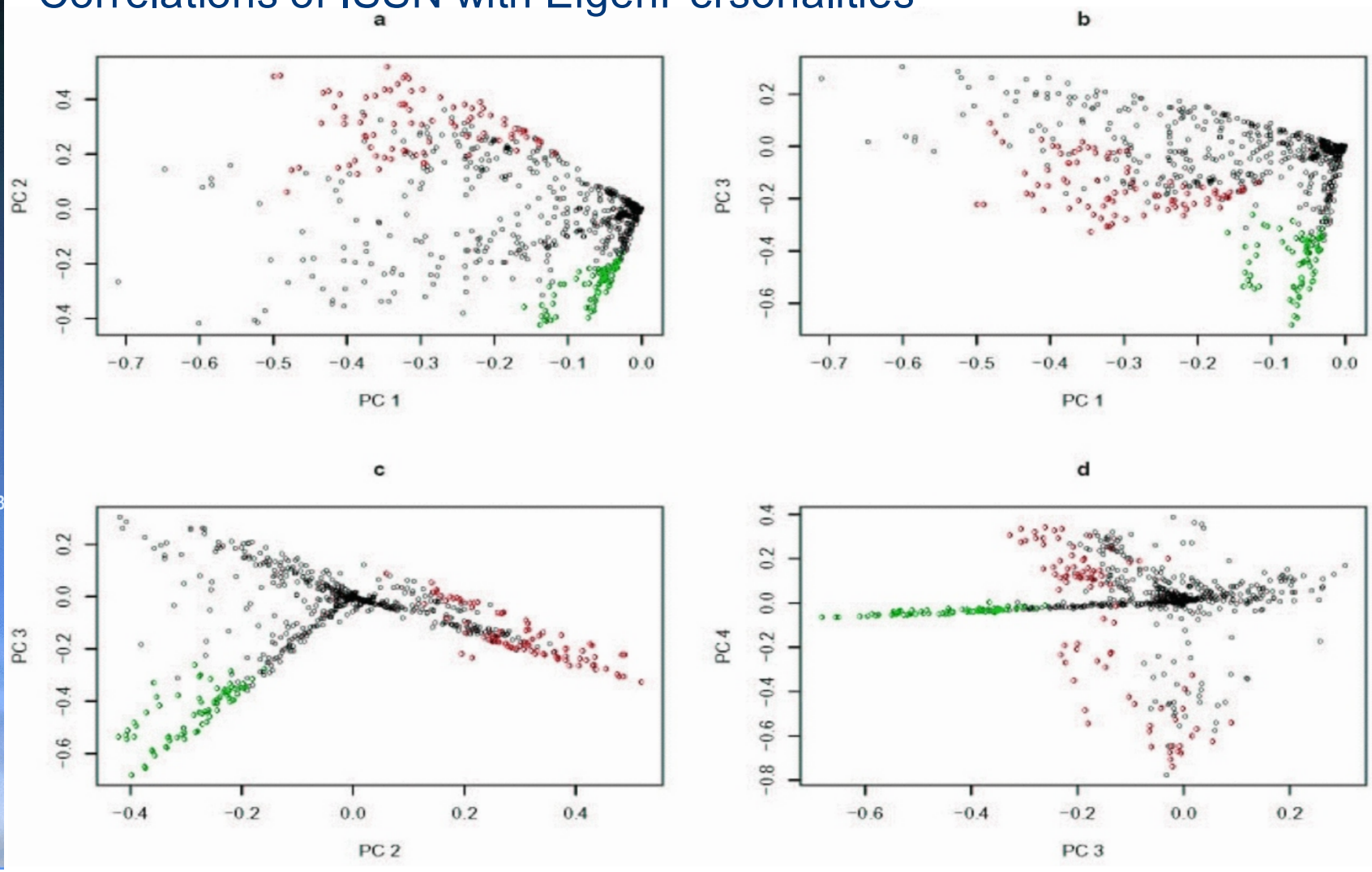


Luis Rocha
2003

rocha@lanl.gov

SVD of Personality ISSN Relation

Correlations of ISSN with EigenPersonalities



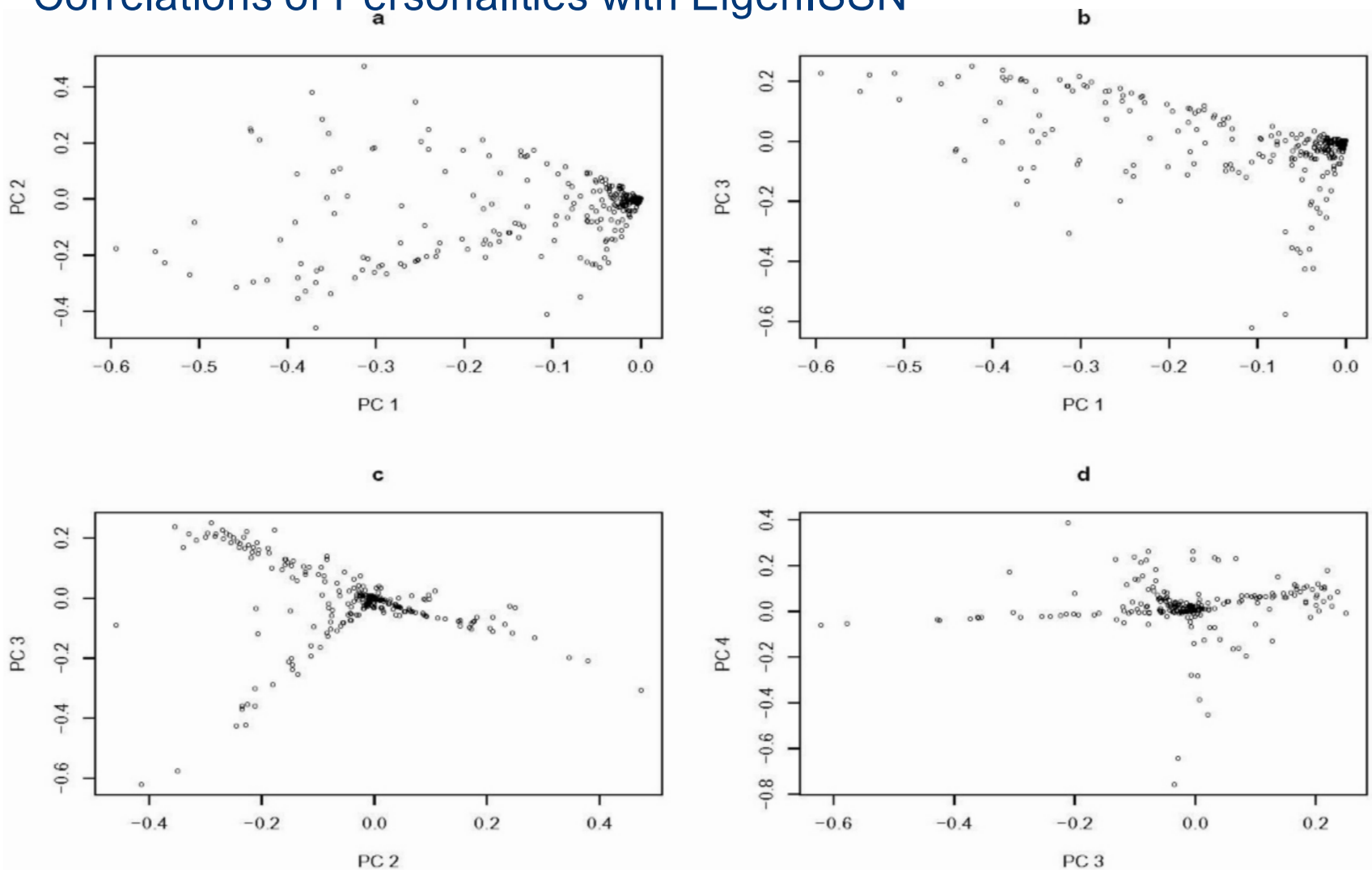
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

SVD of Personality ISSN Relation

Correlations of Personalities with EigenISSN

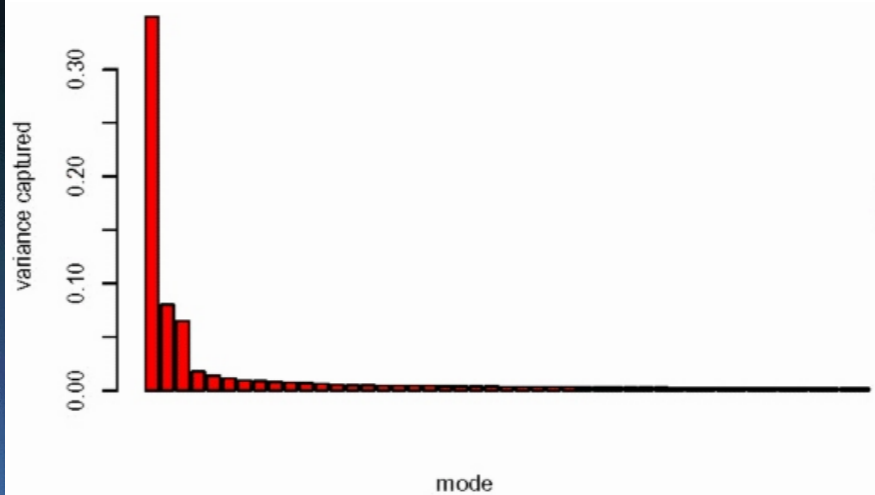


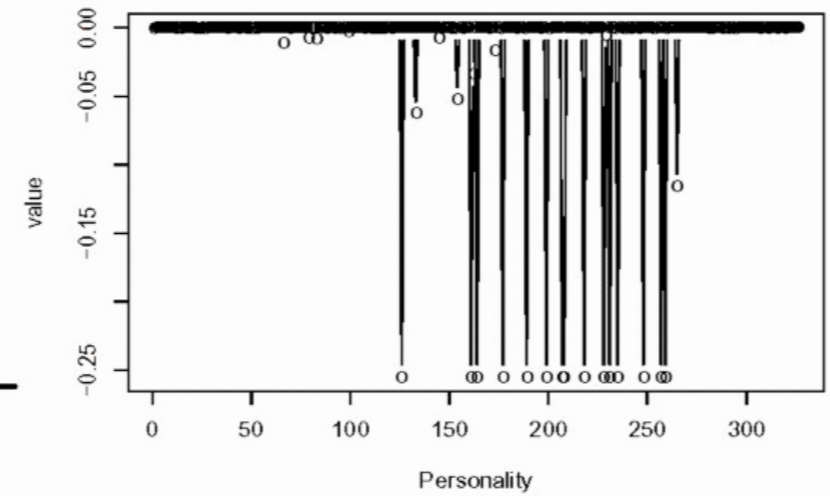
Luis Rocha
2003

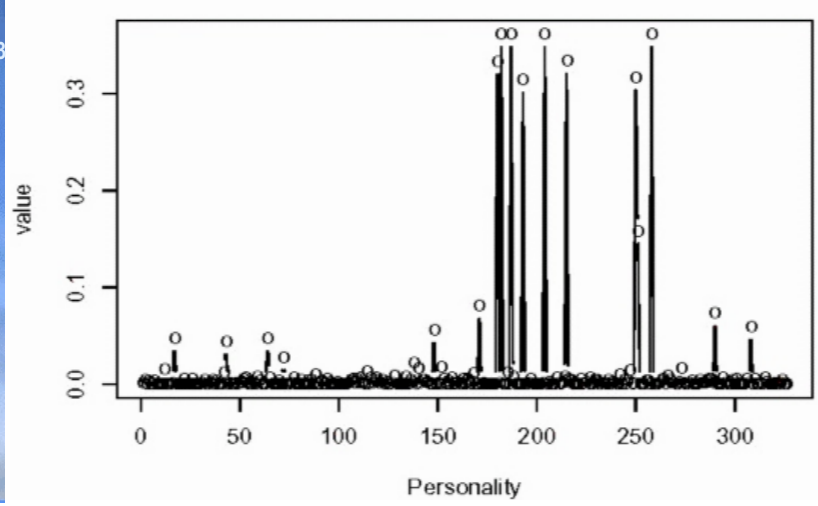
Eigen Analysis of MyLibrary

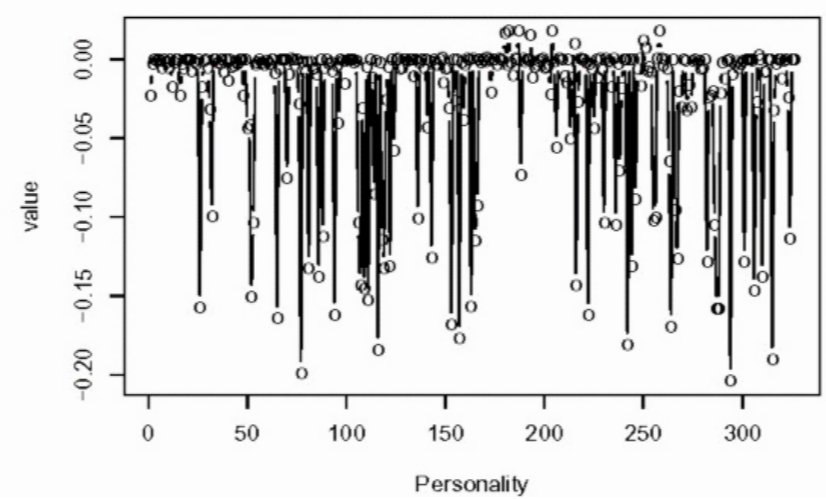
PCA of PIP_4

rocha@lanl.gov





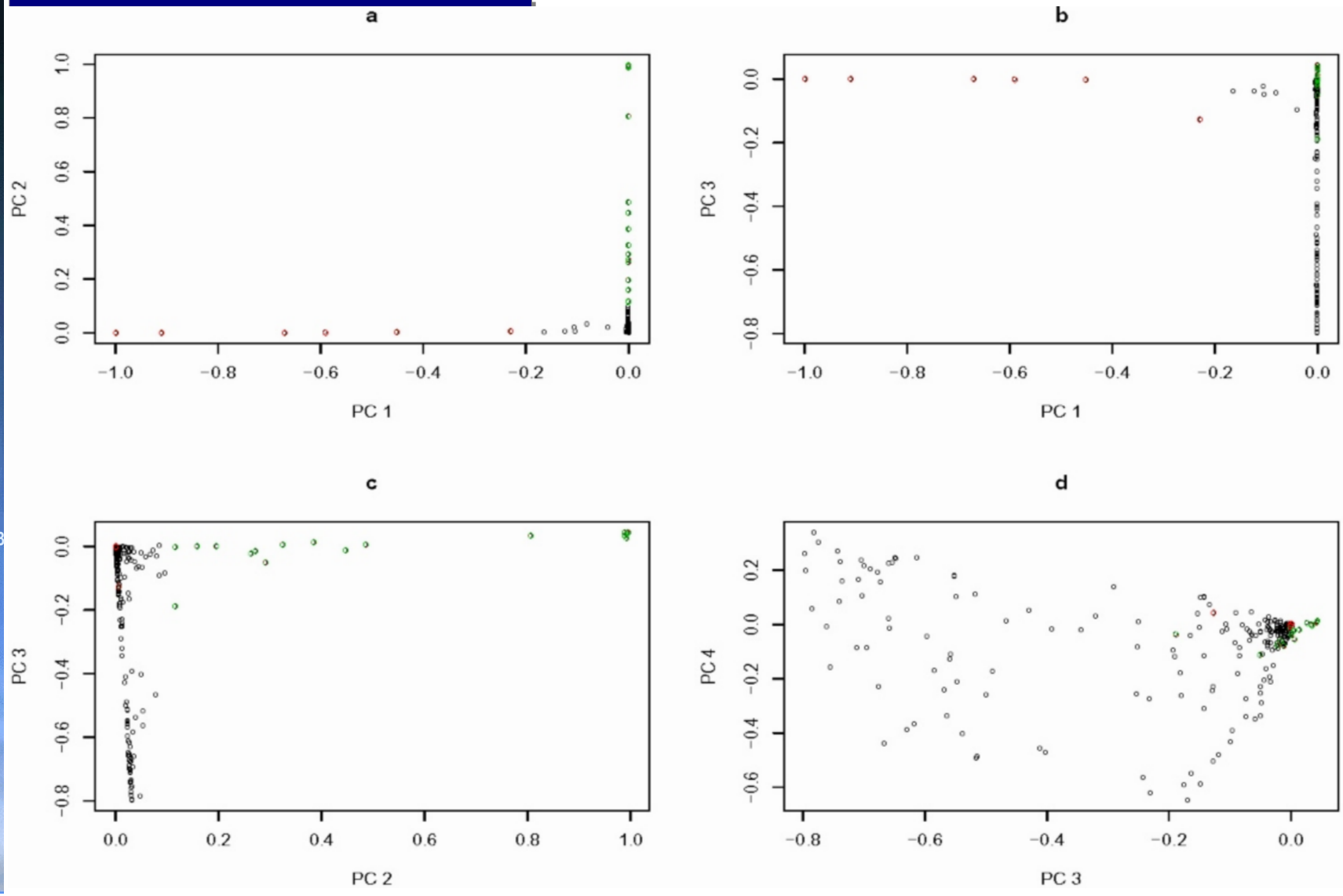




Luis Rocha
2003

PCA of PIP_4

rocha@lanl.gov



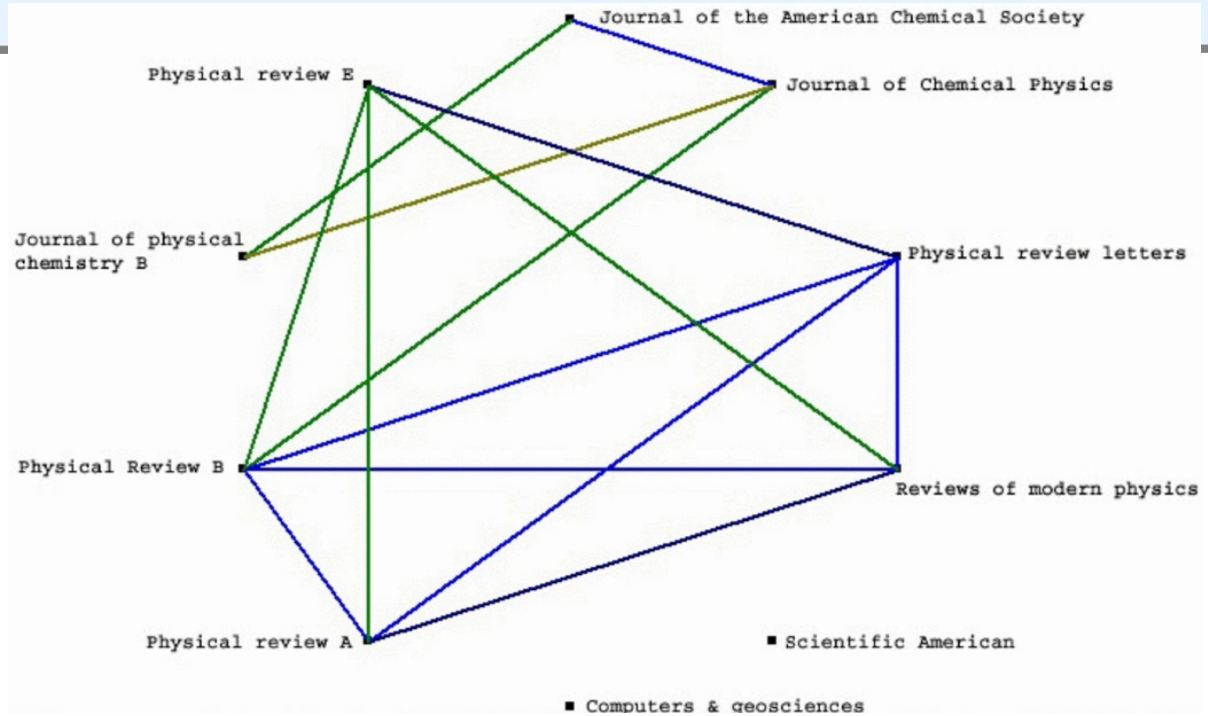
Luis Rocha
2003

rocha@lanl.gov

What about Transitivity?

Things indirectly related

- Proximity from co-occurence
  - Instead of frequency
  - Is Symmetric
- Is there relevance information on transitivity?
  - How far to go?



rocha@lanl.gov

Distance from Proximity

Semi-metric Behavior

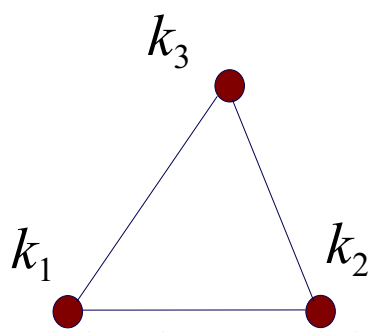
$$kdp\ k_i, k_j \quad \frac{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}}{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}} \quad \frac{N_{\cap}\ k_i, k_j}{N_{\cup}\ k_i, k_j}$$

(Keyword Document Proximity)

$$d_{kdp}\ k_i, k_j \quad \frac{1}{kdp\ k_i, k_j} - 1$$

(Keyword Document Distance)

d is a distance function because it is a nonnegative, symmetric, real-valued function such that $d(k, k) = 0$

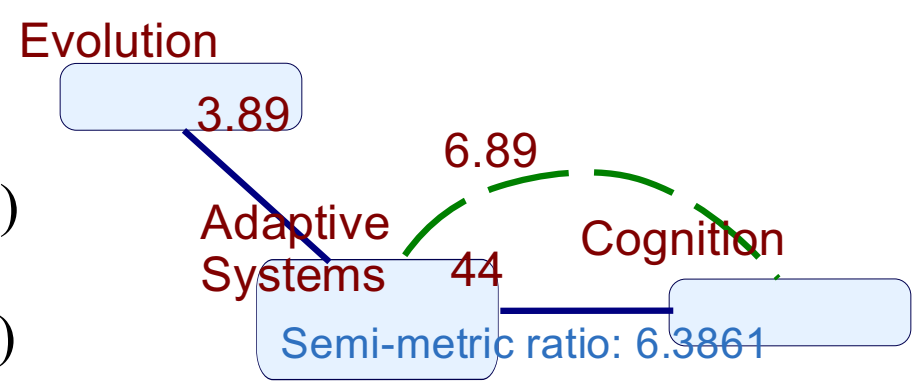


$$d(k_1, k_2) \leq d(k_1, k_3) + d(k_3, k_2)$$

Metric

$$d(k_1, k_2) > d(k_1, k_3) + d(k_3, k_2)$$

Semi-metric



rocha@lanl.gov

Distance from Proximity

Generic Case

$$d_X(x_i, x_j) = \frac{1}{XYP(x_i, x_j)} - 1; \quad d_Y(y_i, y_j) = \frac{1}{YXP(y_i, y_j)} - 1$$

$$XYP(x_i, x_j) = \frac{\sum_{k=1}^{m} (r_{i,k} \wedge r_{j,k})}{\sum_{k=1}^{m} (r_{i,k} \vee r_{j,k})}; \quad YXP(y_i, y_j) = \frac{\sum_{k=1}^{n} (r_{k,i} \wedge r_{k,j})}{\sum_{k=1}^{n} (r_{k,i} \vee r_{k,j})}$$

Distance from a Proximity Graph is semi-Metric
Distance from a Similarity Graph is Metric

rocha@lanl.gov

Measuring Semi-Metric Behavior

Semi-metric Measures

- **Semi-metric ratio**
  - Absolute measure of indirect distance reduction

$$b(x_i, x_j) = \frac{\overline{d_{x_i}}}{d_{shortest}(x_i, x_j)}$$

- **Relative Semi-metric ratio**
  - Distance reduction against maximum contraction

$$rs(x_i, x_j) = \frac{d_{direct}(x_i, x_j) - d_{shortest}(x_i, x_j)}{d_{\max} - d_{\min}}$$

- **Below Average Ratio**
  - Captures semi-metric distance reductions which contract to below the average distance for a given node. Captures some of the cases of initial ∞ distance

$$x_i, x_j) = \frac{d_{direct}(x_i, x_j)}{d_{shortest}(x_i, x_j)}$$

rocha@lanl.gov

Semi-metric Analysis

Applied to Many Data Sets

- Keywords extracted from books cited in a dissertation
  - Identified novel theme associations
- ARP Database
  - 3 million scientific articles published 1996-2000
  - Identified trends
- Web Sites, Web logs, Word Norm, Random Graphs
  - Denotes Latent Associations

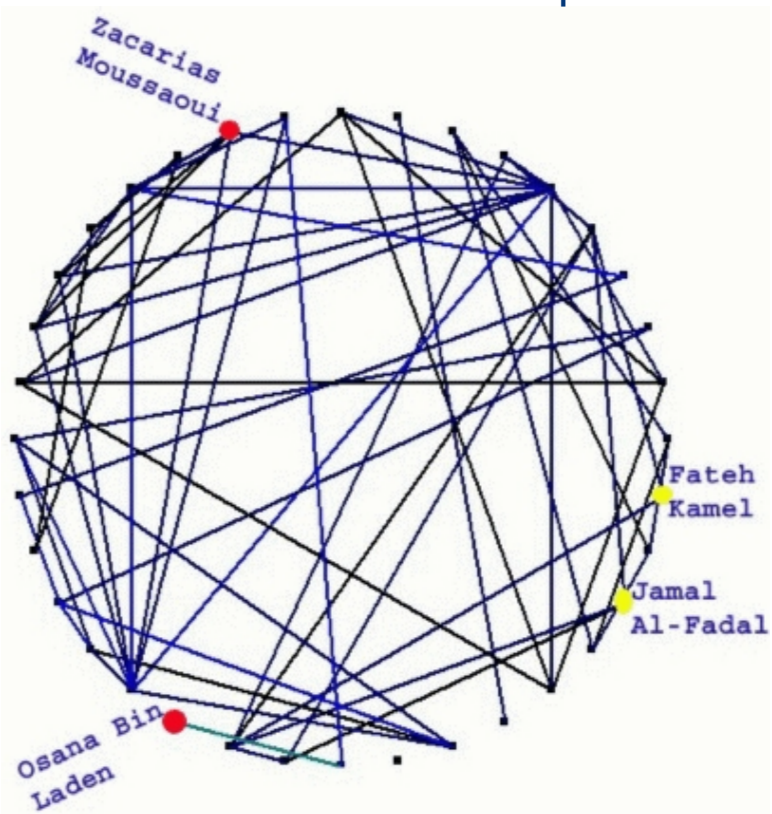
Rocha, Luis M. [2002]. "Semi-metric Behavior in Document Networks and its Application to Recommendation Systems". In: *Soft Computing Agents: A New Perspective for Dynamic Information Systems.* V. Loia (Ed.) International Series Frontiers in Artificial Intelligence and Applications. IOS Press, pp. 137-163.

Rocha, Luis M. [2002]. "Combination of Evidence in Recommendation Systems Characterized by Distance Functions". In: *Proceedings of the 2002 World Congress on Computational Intelligence: FUZZ-IEEE'02*. Honolulu, Hawaii, May 2002. IEEE Press, pp. 203-208.

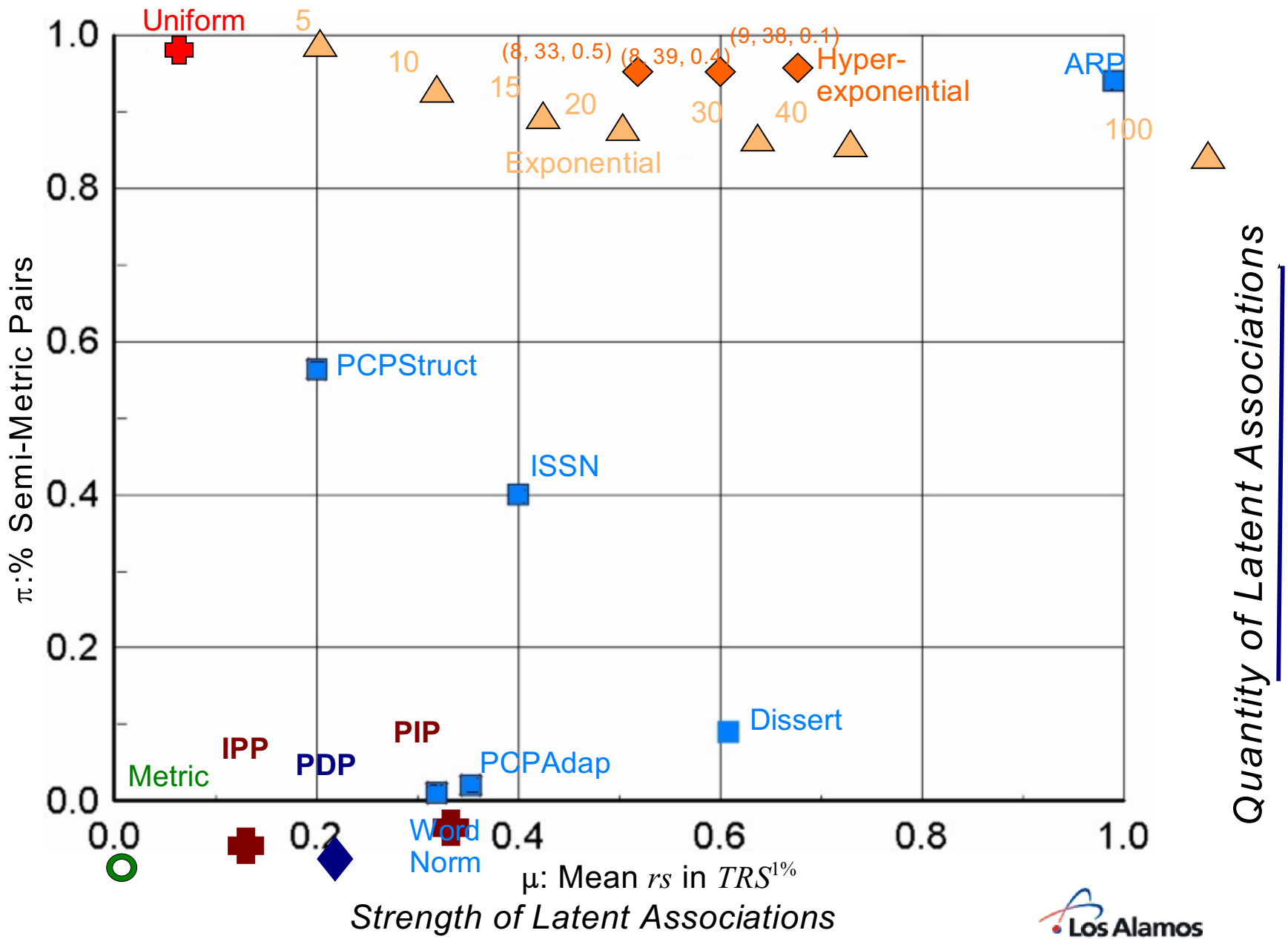
rocha@lanl.gov

Identification of Latent Associations

Terrorist Network Example



- Pairs with larger semi-metric behavior denote a *latent association*
  - Not grounded on direct evidence provided by the relation R, but rather implied by the overall network of associations in this relation.
  - Meaning depends on the semantics of the application
    - In graphs of keyword co-occurrence in documents: associated with novelty and can be used to identify trends.
    - In terrorist networks it may identify pairs of people, groups, etc. for which we do not have direct evidence, in the available documents, that a real association exists, but who could easily be indirectly associated.
  - In recommendation system for journals now at LANL



Los Alamos NATIONAL LABORATORY

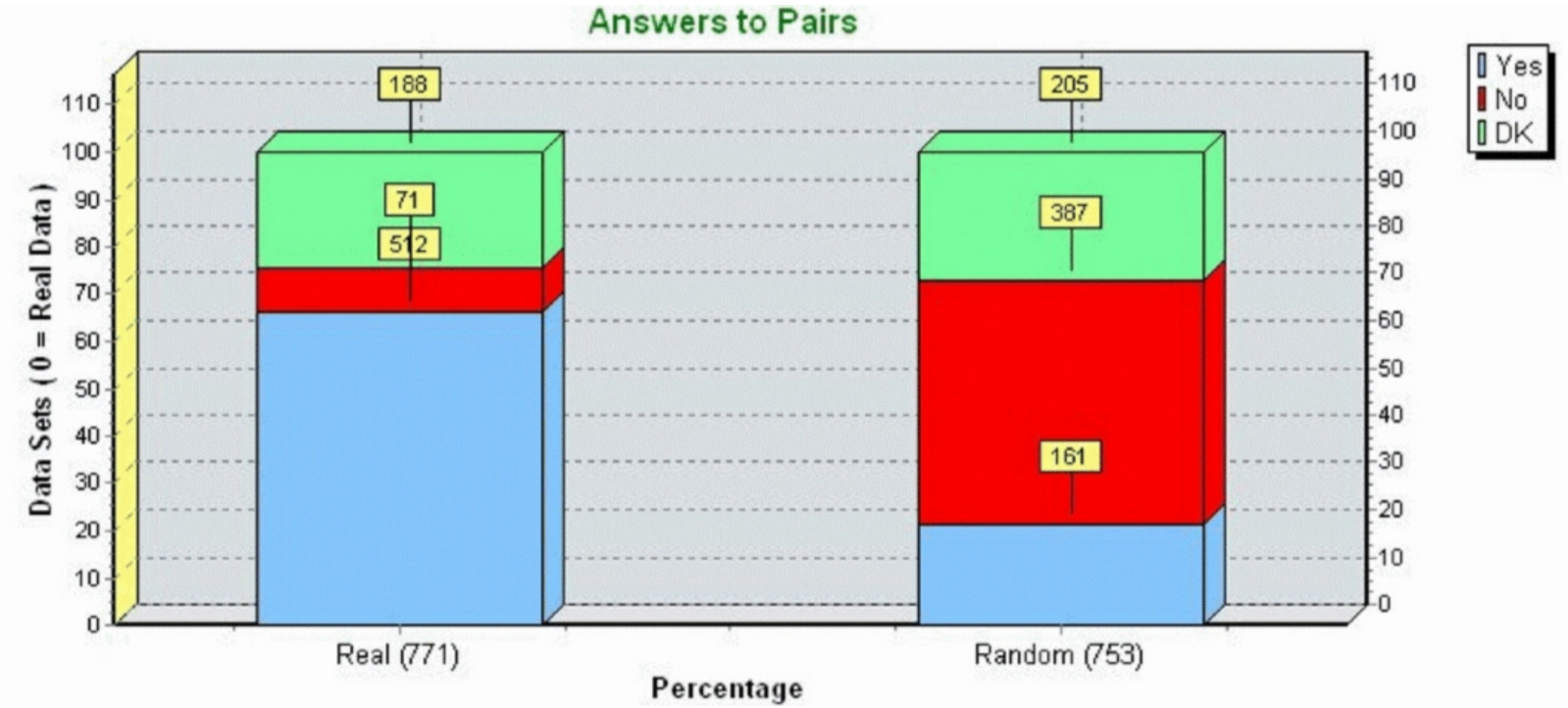
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

User Test

Population from STB-RL and CCS-3



- 50% Pairs of ISSN with high *rs* and *b*
- 50% Random ISSN Pairs
- 1524 answers

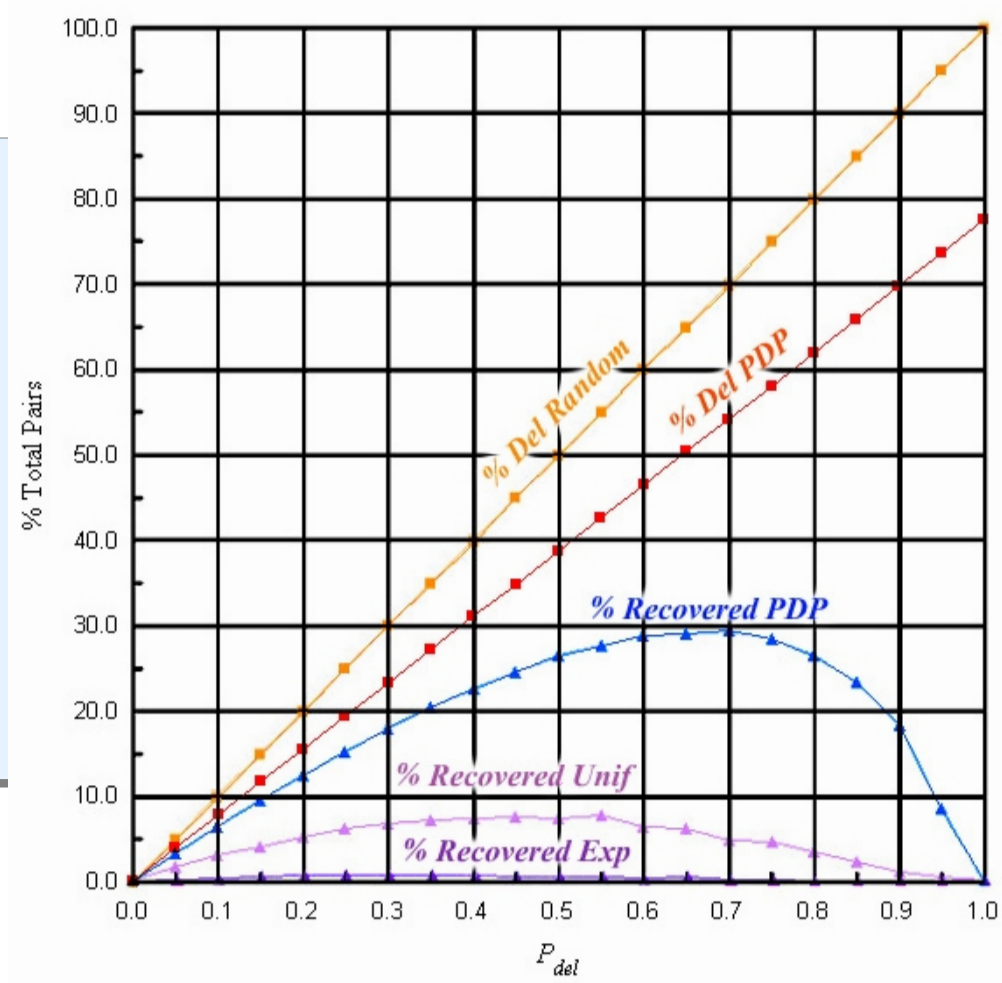
Luis Rocha
2003

rocha@lanl.gov

Detecting Incomplete Knowledge

Random Deletion Experiments

- Perfect Knowledge
  - Transitive Closure of real graph
  - Metric Distance Graph
- Incomplete Knowledge
  - Each positive association is deleted with probability p_{del}
  - 100 graphs for each value of p_{del}



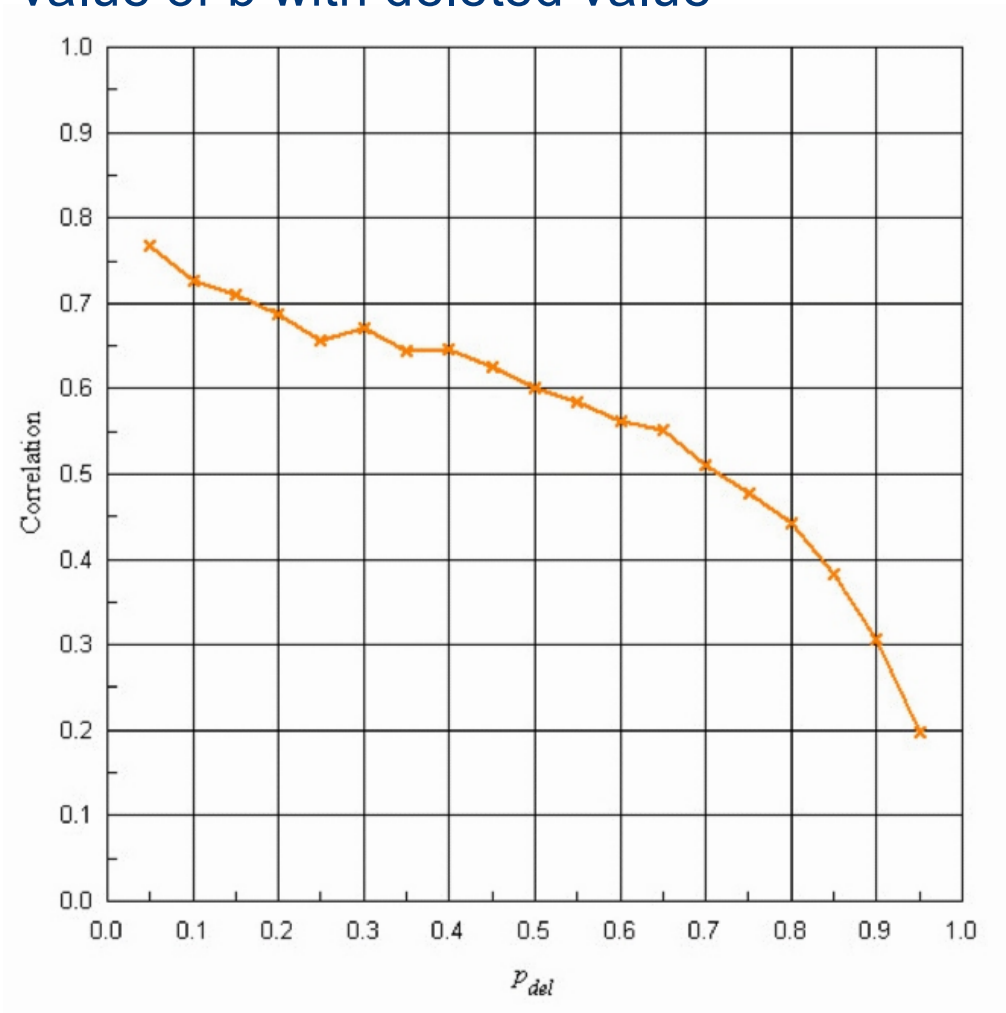
Full Deletion

Recovery via parameter *b*

rocha@lanl.gov

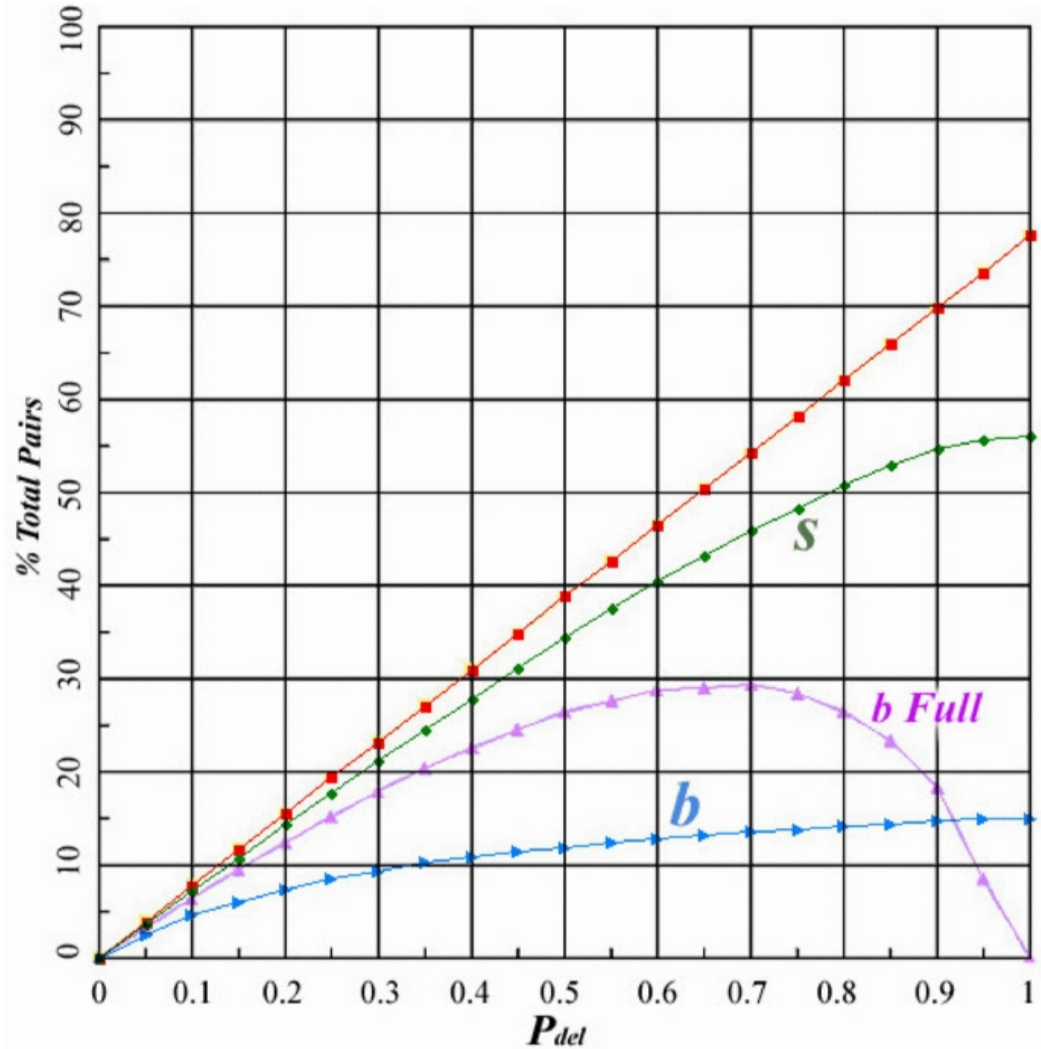
Full Deletion

Correlation of value of b with deleted value



rocha@lanl.gov

Partial Deletion

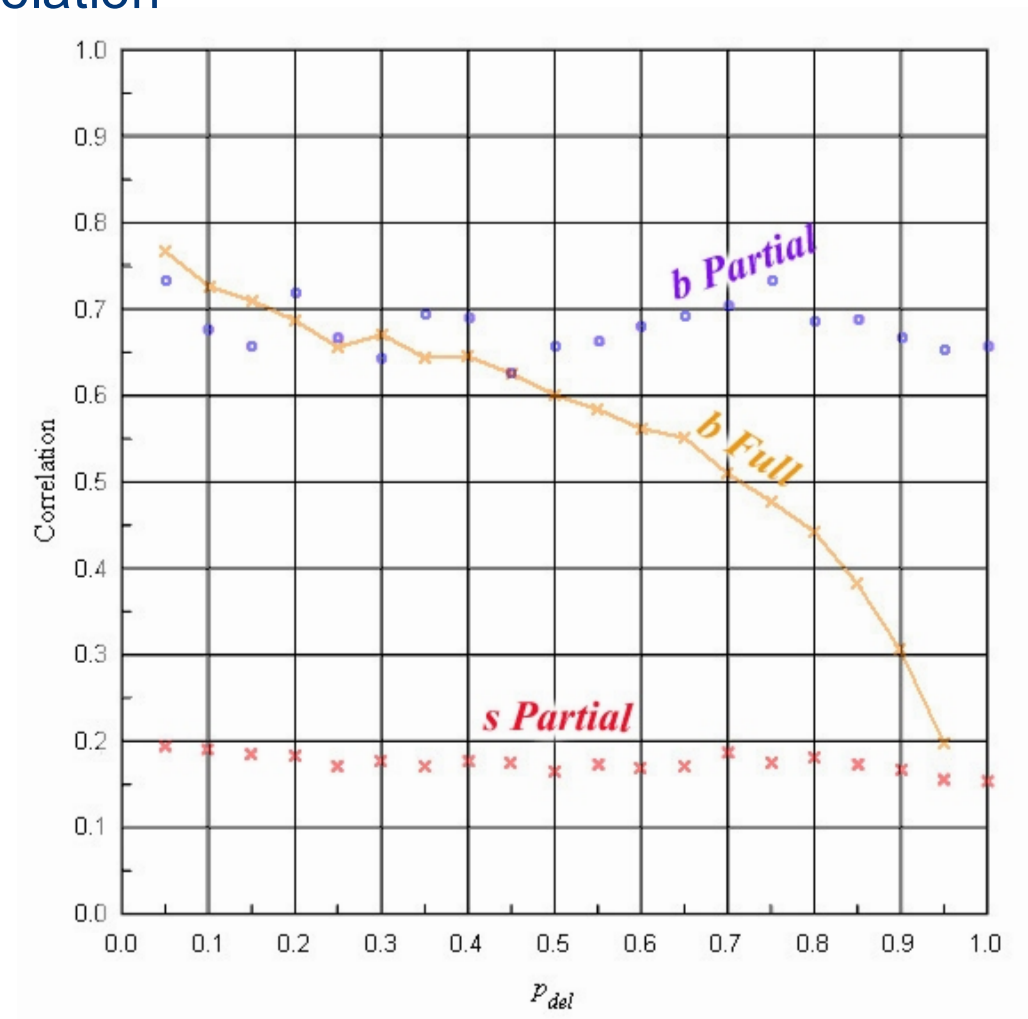


Luis Rocha
2003

rocha@lanl.gov

Partial Deletion

Parameter Correlation



Luis Rocha
2003

rocha@lanl.gov

Semi-Metric Recommendations

IPP_3: parameters *rs* and *b*

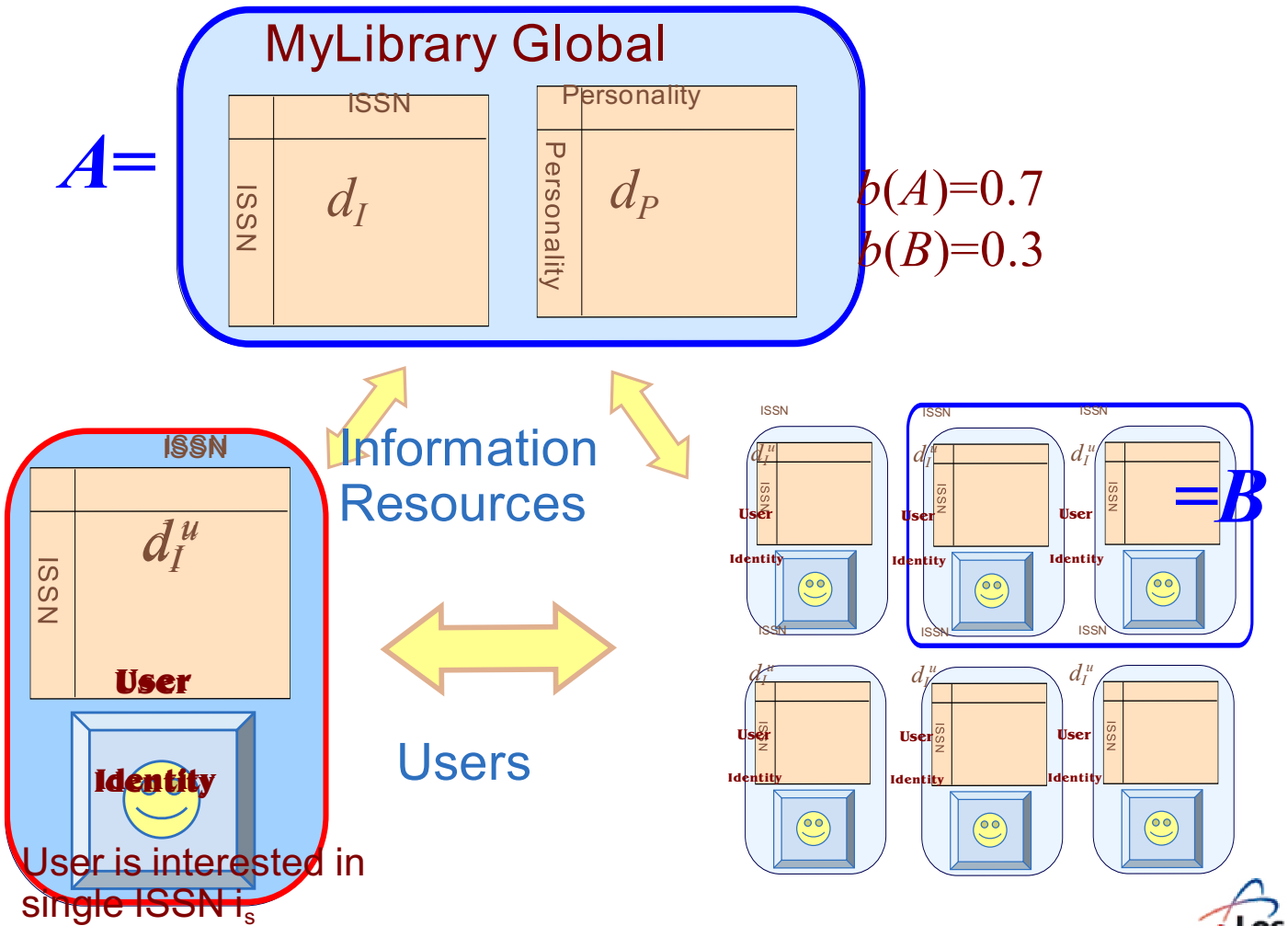
0020-1669--Inorganic chemistry 0031-9007--Physical review letters
0031-9007--Physical review letters 0743-7463--Langmuir
0003-2700--Analytical chemistry 0031-9007--Physical review letters
0096-3003--Applied mathematics and computation 0031-9007--Physical review letters
0031-9007--Physical review letters 0022-3115--Journal of nuclear materials
1049-3301--ACM transactions on modeling and computer simulation 0031-9007--Physical review letters
1364-548X--Chemical communications 0031-9007--Physical review letters
1064-8275--SIAM journal on scientific computing 0031-9007--Physical review letters
0965-5425--Computational mathematics and mathematical physics 0031-9007--Physical review letters
0031-9007--Physical review letters 1359-6454--Acta materialia
0003-7028--Applied spectroscopy 0031-9007--Physical review letters
0031-9007--Physical review letters 0022-2461--Journal of materials science
0031-9007--Physical review letters 1359-6462--Scripta materialia
0031-9007--Physical review letters 0022-4596--Journal of solid state chemistry
0031-9007--Physical review letters 0021-8898--Journal of applied crystallography

1097-6256--Natureneuroscience 1065-9471--Human Brain MApping
1097-6256--Natureneuroscience 0278-0062--IEEE transactions on medical imaging
1097-6256--Natureneuroscience 1053-8119--NeuroImage
1063-7796--Physics of particles and nuclei 0218-3013--International journal of modern physics E Nuclear physics
1053-8119--NeuroImage 1065-9471--Human Brain MApping
0031-9007--Physical review letters 0743-7463--Langmuir
0031-9007--Physical review letters 0020-1669--Inorganic chemistry
0031-9007--Physical review letters 0141-1594--Phase transitions
0031-9007--Physical review letters 0928-1045--Journal of computer aided materials design
0031-9007--Physical review letters 0042-207X--Vacuum
1097-6256--Natureneuroscience 0031-9155--Physics in medicine & biology
1097-6256--Natureneuroscience 0096-3518--IEEE transactions on acoustics speech and signal processing
1097-6256--Natureneuroscience 0740-7467--IEEE ASSP magazine
1097-6256--Natureneuroscience 1070-9908--IEEE signal processing letters
0022-5355--Journal of vacuum science and technology 0734-2101--Journal of vacuum science & technology A Vacuum surfaces and films

TalkMine For MyLIbrary

Distributed Conversation

rocha@lanl.gov



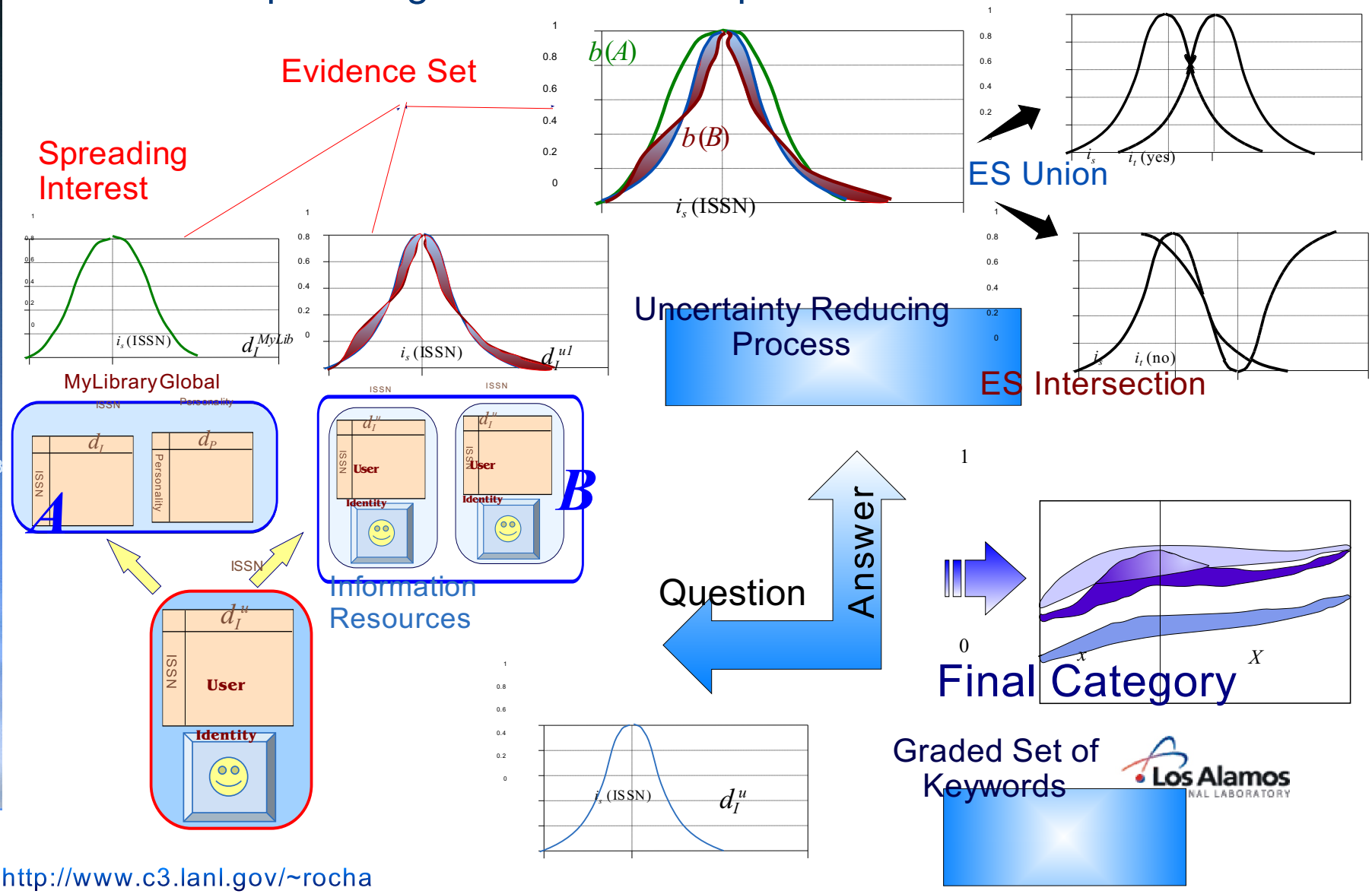
Luis Rocha
2003

http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

TalkMine: Categorization

Interactive Spreading Interest on Multiple Resources



Luis Rocha
2003

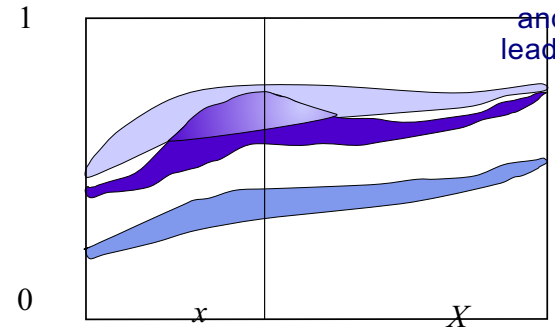
http://www.c3.lanl.gov/~rocha

rocha@lanl.gov

Evolution of New Keywords

No Categories are stored as such, but rather constructed in conversation between agents and information resources, leading to agent identification

Short-Term Categorization



$$N^{A}_{t+1}(i_s \cap i_t) = N$$

$$N^{A}_{t+1}(i_s) = N^{A}_{t}(i_s$$

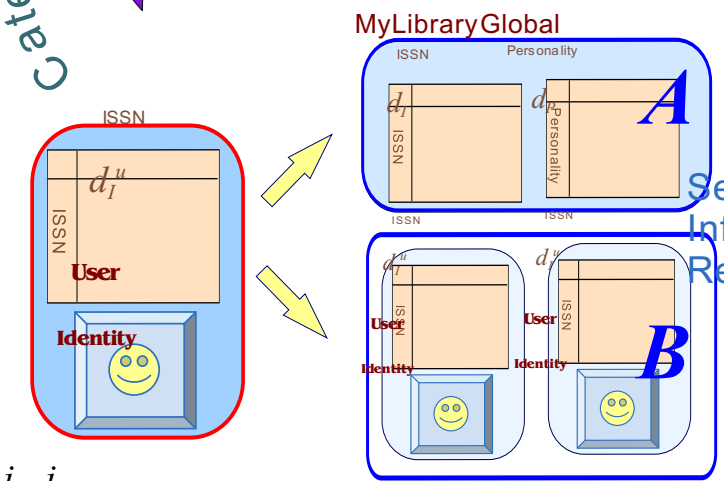
Categorization

Final Category

Adaptation

Long-Term Memory

New ISSN are recognized by Information Resources



Set of Information Resources

$$ipp\ i_s, i_t \quad \frac{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}}{\sum_{k=1}^{m} a_{i,k}\ a_{j,k}} \quad \frac{N_{\cap}\ i_s, i_t}{N_{\cup}\ i_s, i_t}$$

(ISSN Personality Proximity)

Luis Rocha
2003

http://www.c3.lanl.gov/~rocha