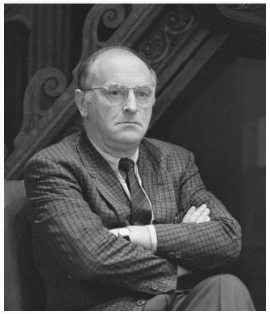
University of Strathclyde

Communicability and Geometry of Complex Networks

Ернесто Естрада

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, SCOTLAND

ernesto.estrada@strath.ac.uk
www.estradalab.org
@eestradalab



“The true beauty is the beauty at low temperatures”

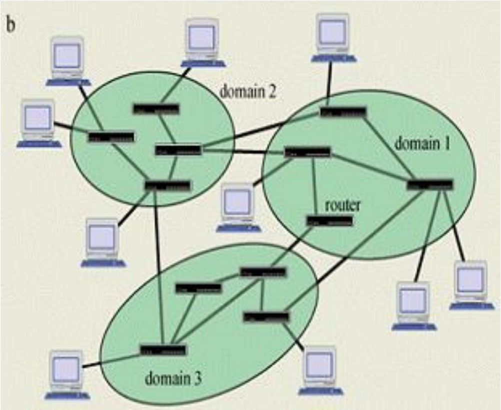
Ио́сиф Алекса́ндрович Бро́дский

Outline

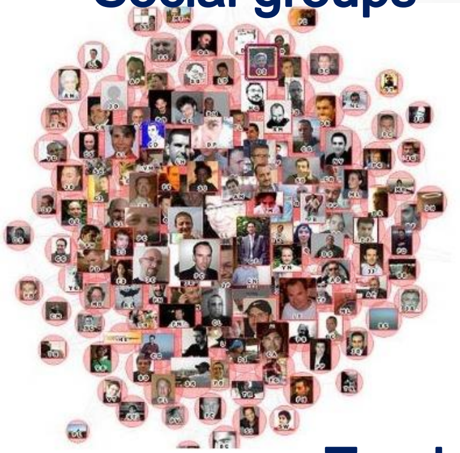
- ✓ Motivating communicability
- ✓ Emergence of a geometry
- ✓ Communicability distance
- ✓ Communicability angles and spatial efficiency
- ✓ Communicability angles and diffusion

Networks. Informally…

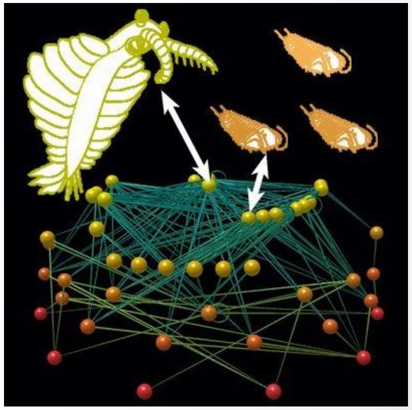
Infrastructures



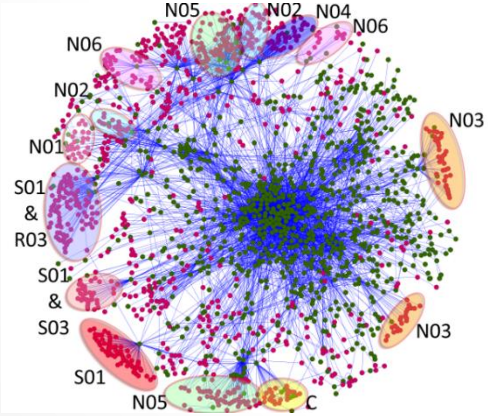
Social groups



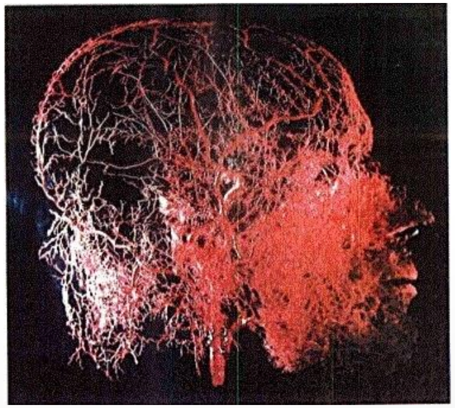
Trophic relations



Molecular interactions

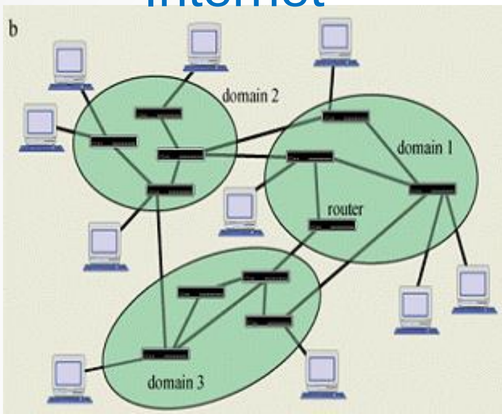


Anatomical systems

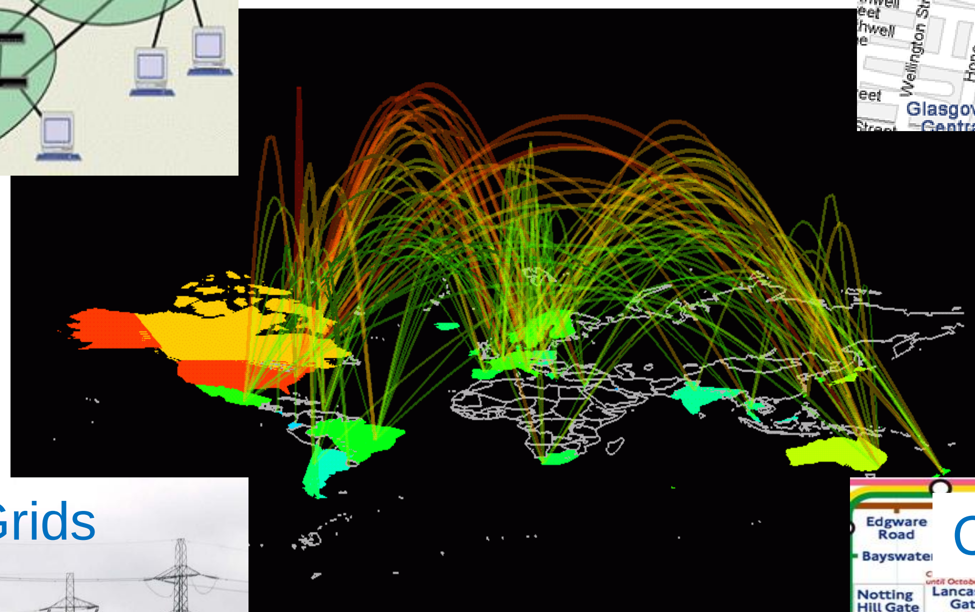


Networks. Informally…

Internet



City Networks

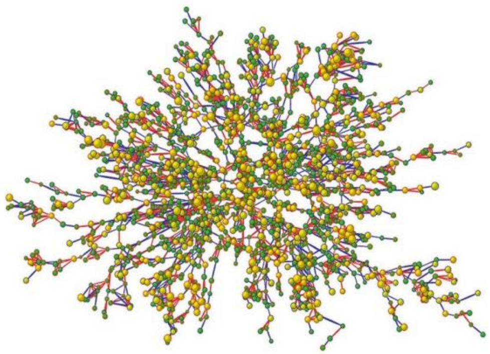


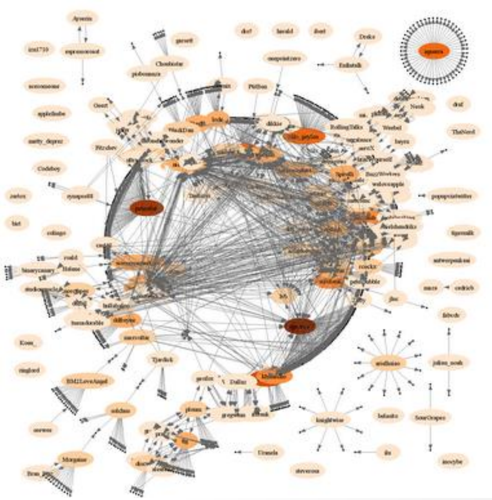
Power Grids

Communication

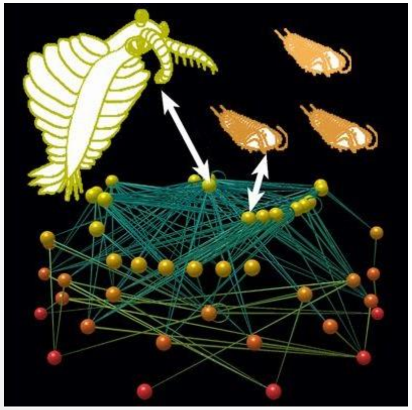
Networks. Informally…

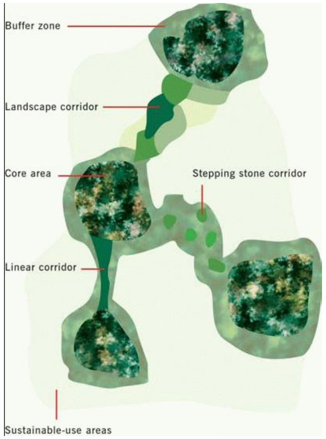
Social Networks

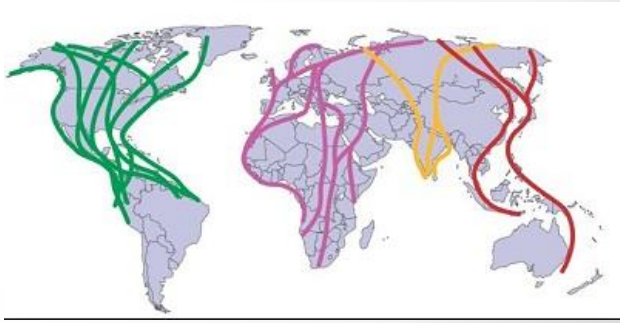




Ecological Networks

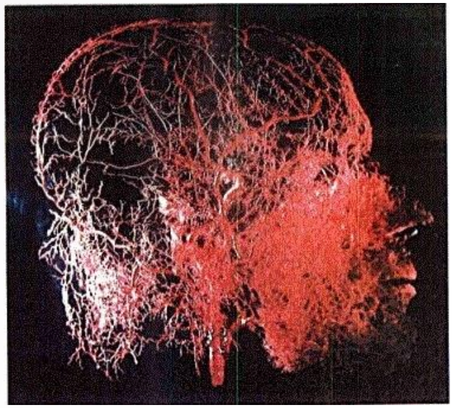


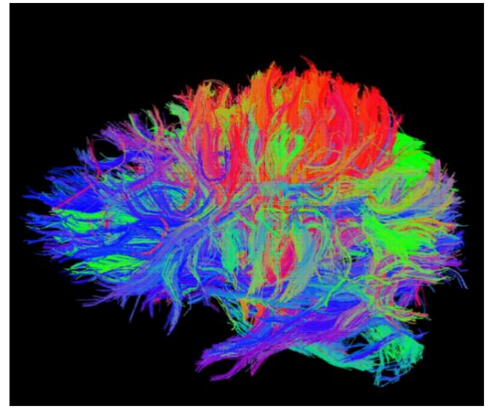


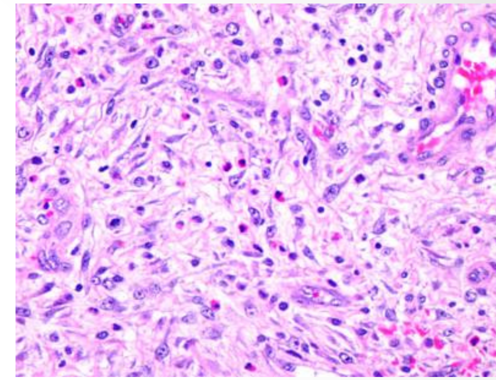


Networks. Informally…

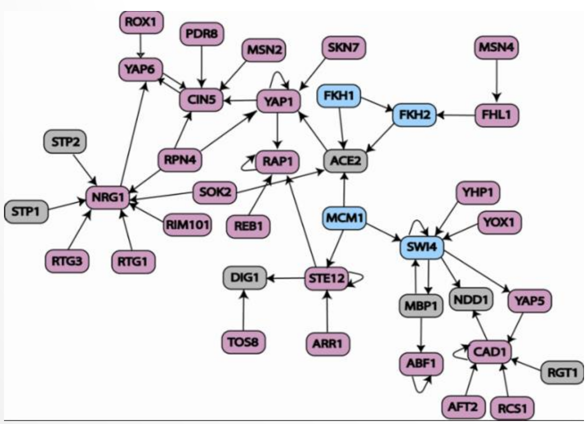
Anatomical Networks

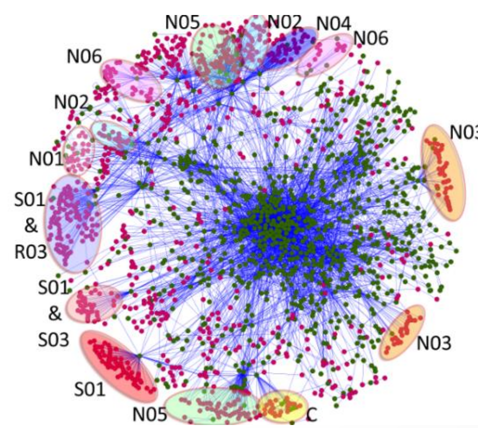


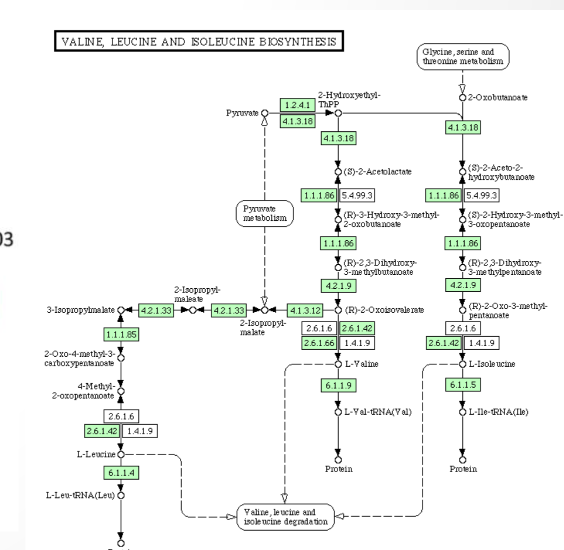




Molecular Networks

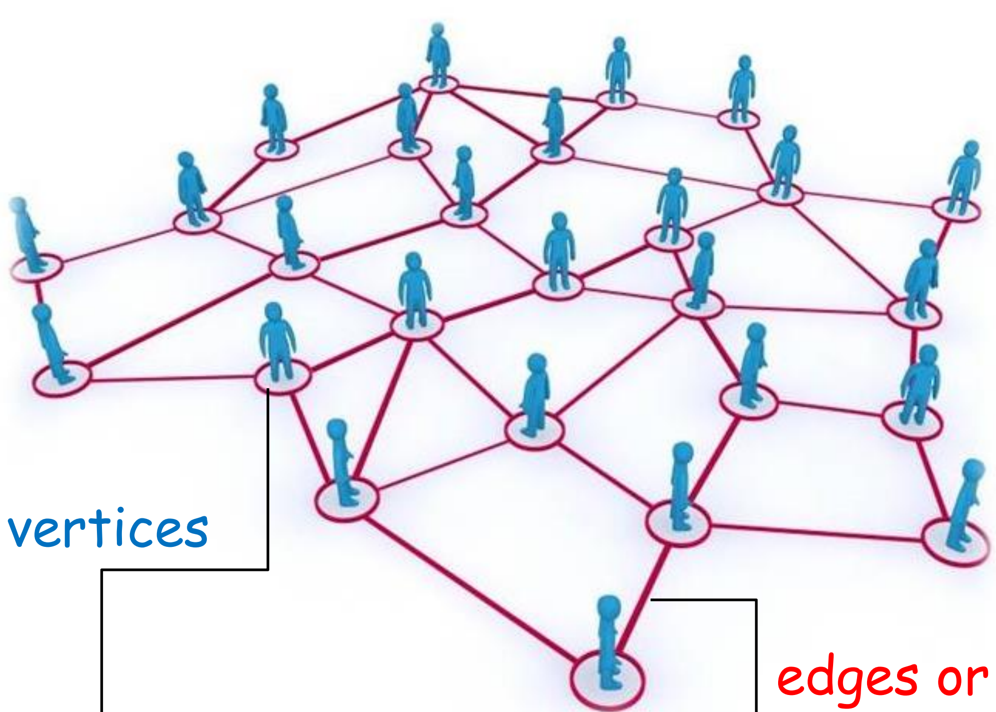






Networks. Formally…

$$G = (V, E)$$



nodes or vertices

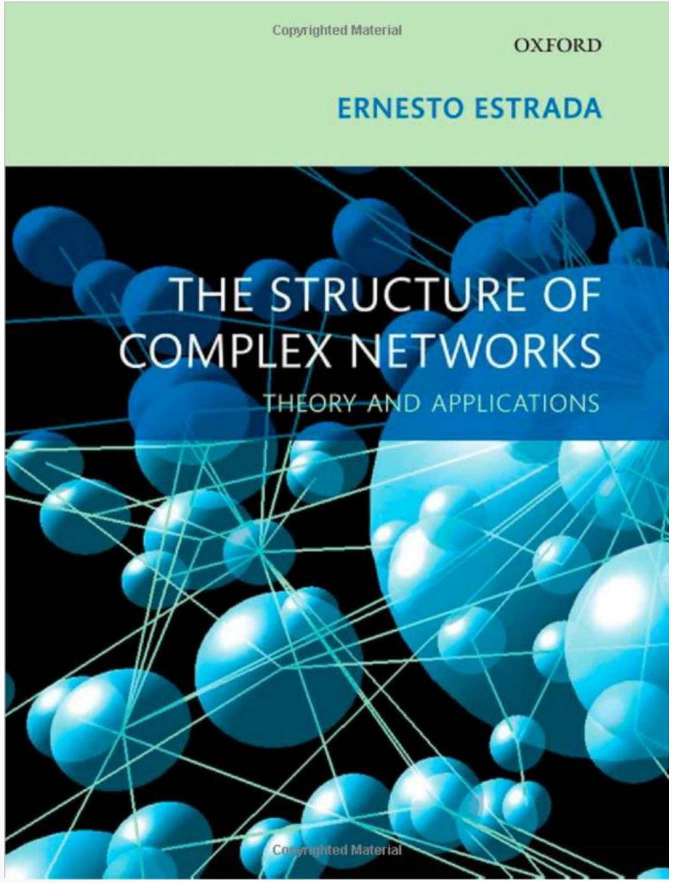
edges or links

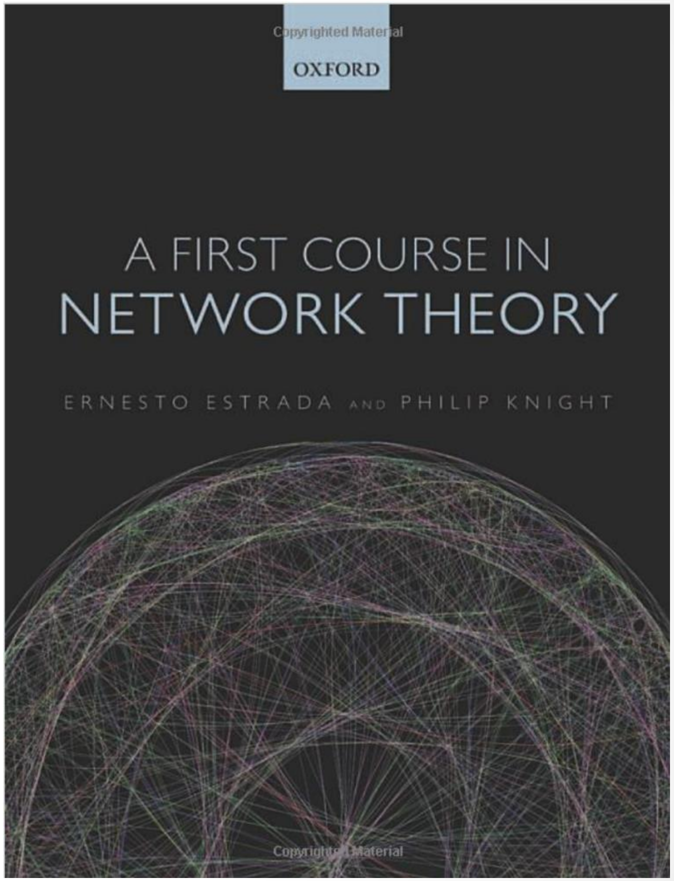
$$V = \{v_1, \cdots, v_n\}$$

$$E \subseteq V \times V$$

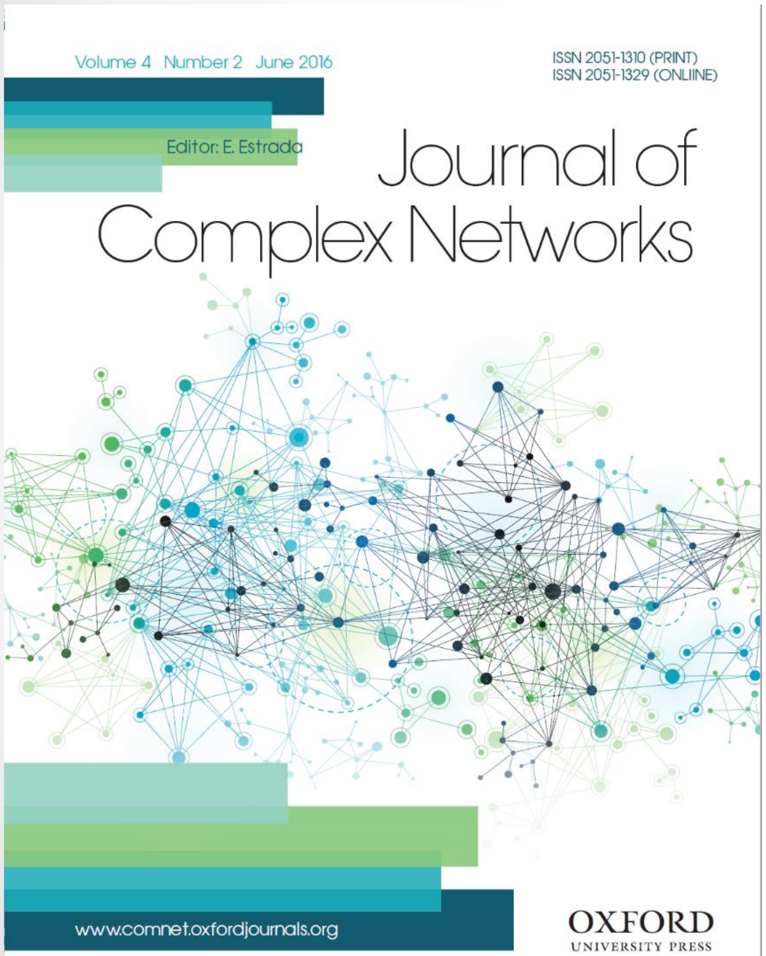
Estrada & Knight: *A First Course on Network Theory,* Oxford Univ. Press, 2015

Useful information





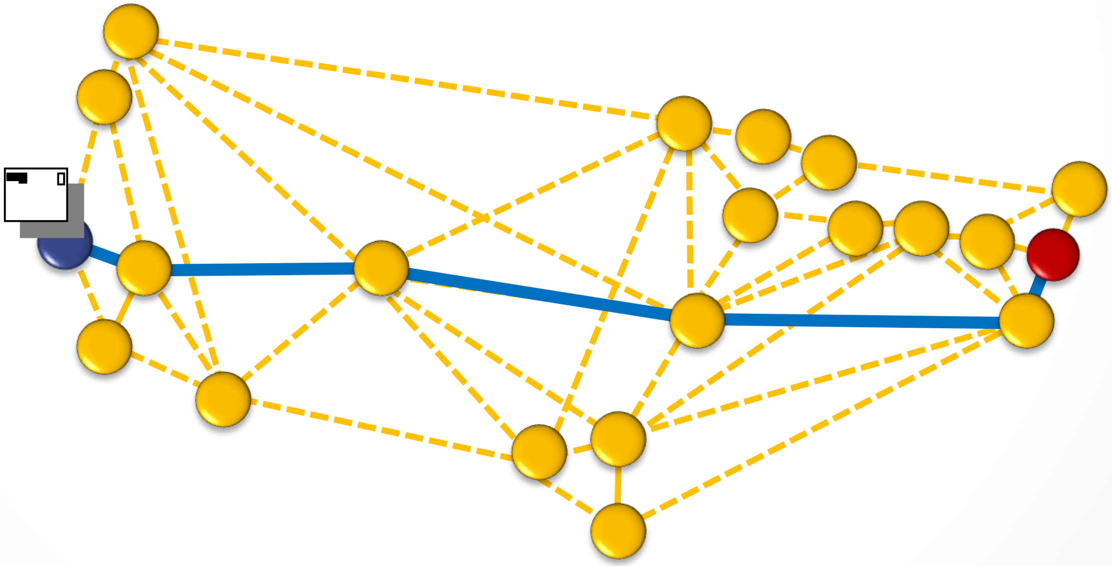
Useful information



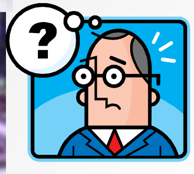
Estimated IF: 4.18 (2017)

Shortest-path (SP) communication

It is rational to think that the communication between two nodes in a network occurs via the shortest path connecting them.

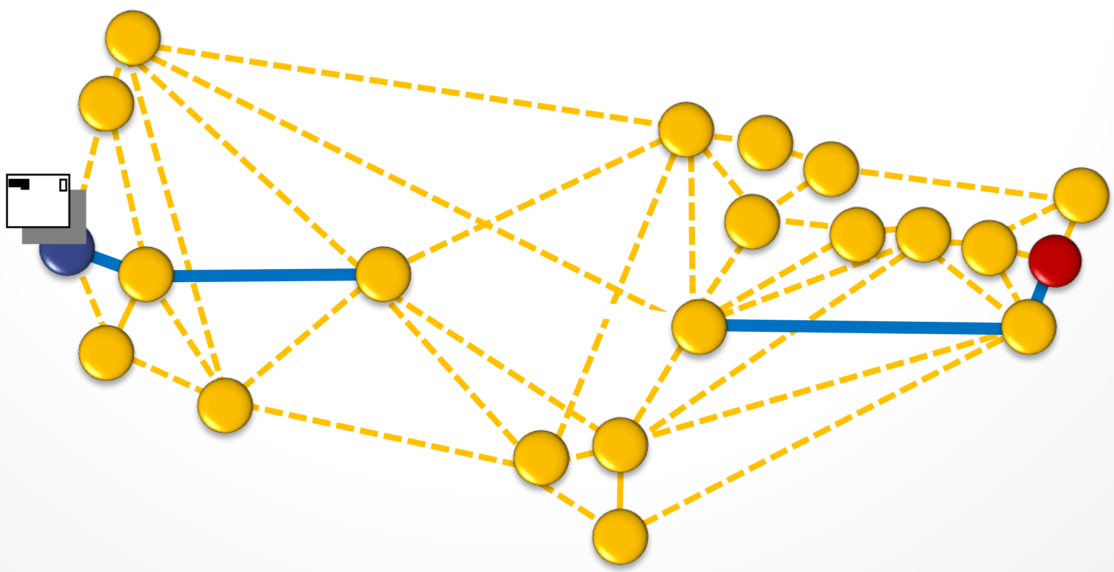


Shortest-path (SP) communication

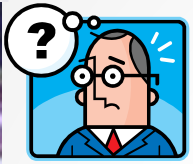


Problem 1) The sender does not know the global structure of the network and she can not know which of the many routes connecting it with the destination is the shortest one.

Problem 2) If the sender knows the shortest path, she does not know *a priori* whether there are damaged edges in it.

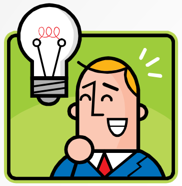


Shortest-path (SP) communication

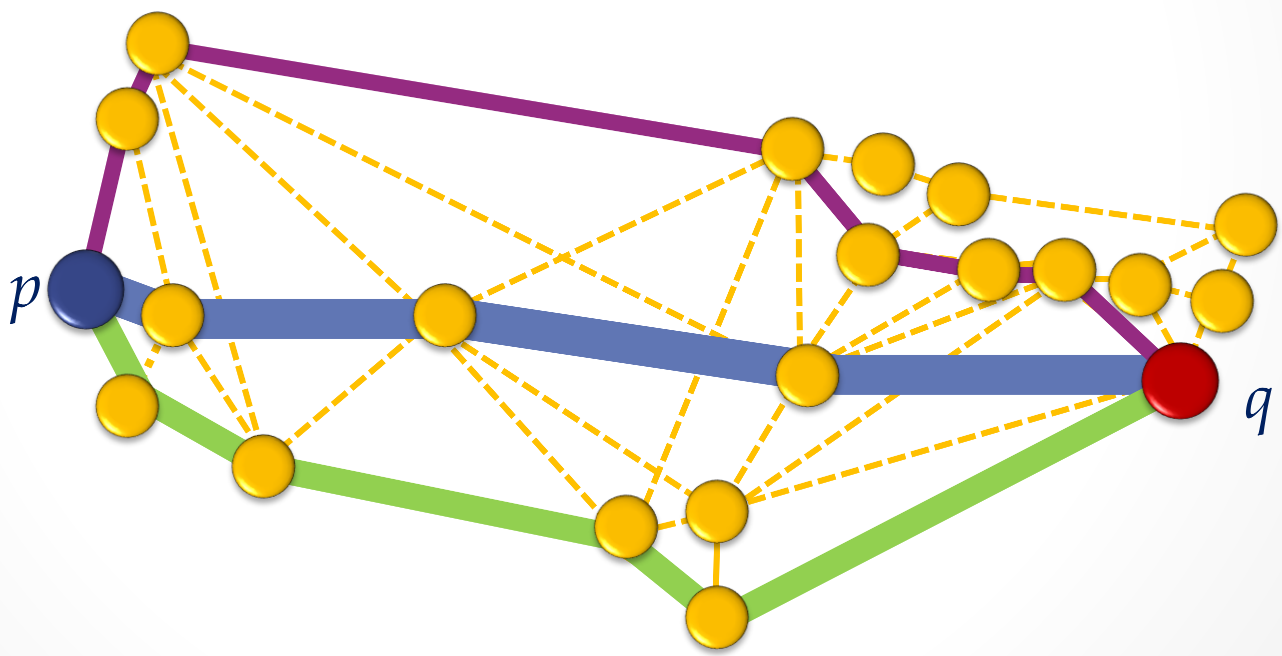


Although in a tree there is a unique shortest path between every pair of nodes, there are no “complex networks” with a tree structure.

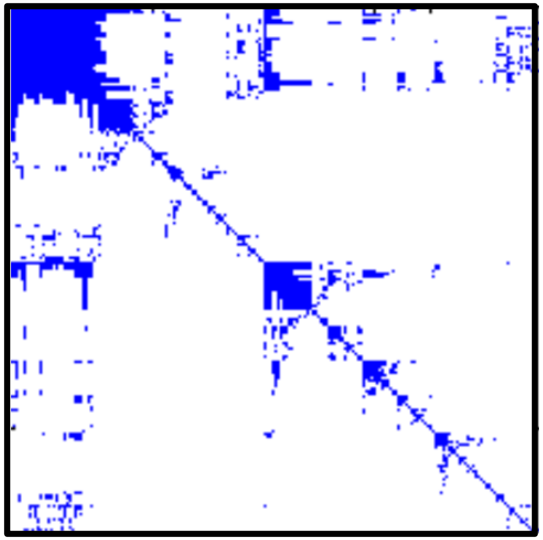
Shortest-path (SP) communication



It seems that the structural redundancy present in complex networks is necessary for guaranteeing the communication among the nodes.



Adjacency Matrix

$\mathbf{A} =$ 

$$A_{ij} = \begin{cases} 1 & (i,j) \in E \\ 0 & (i,j) \notin E \end{cases}$$

Eigenvalues of $\mathbf{A}$: $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n$

Eigenvector of λ_j: $\vec{\varphi}_j = \left[\varphi_j(1), \cdots, \varphi_j(n)\right]^T$

Walks in Graphs

Definition 1: A *walk* of length l, is any sequence of (not necessarily different) nodes $v_1,\ldots, v_l$ such as for each $i=1,\ldots,l$ there is a link from v_i to v_{i+1}.

Theorem 0 (Cvetković). *The number of walks of length l between the nodes p and q in a network is equal to:* $\left(A^l\right)_{pq}$

The Communicability Concept

Definition 2. *The communicability function is given by the total number of walks, weighted in a monotonic decreasing order of their lengths, connecting the vertices p and q in a network G*

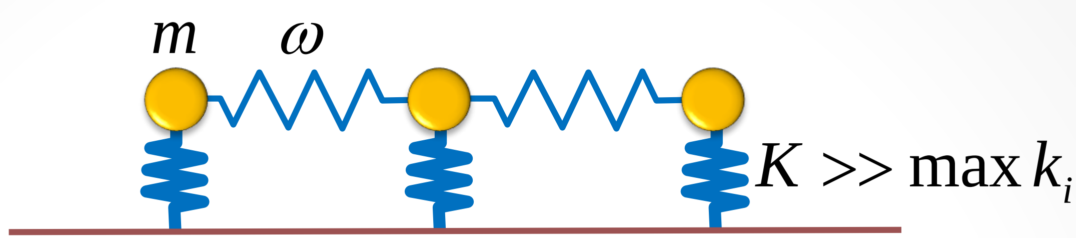
$$G_{pq} = \sum_{k=0}^{\infty} \frac{\left(A^k\right)_{pq}}{k!} = \left(e^A\right)_{pq} = \sum_{j=1}^{n} \psi_{j,p} \psi_{j,q} e^{\lambda_j}$$

where the weighting $(k!)^{-1}$ is selected arbitrarily among the several possibilities.

Estrada & Hatano: *Phys. Rev. E* 77, **2008**, 036111
Estrada, Hatano, Benzi, *Phys. Rep.* 514 **2012**, 89-119

University of Strathclyde

Communicability Function



$$H = \sum_i \hbar\Omega\left(a_i^\dagger a_i + \frac{1}{2}\right) - \frac{\hbar\omega^2}{4\Omega}\sum_{i,j}\left(a_i^\dagger + a_i\right)A_{ij}\left(a_j^\dagger + a_j\right)$$

Theorem 1. *The communicability function corresponds (apart from physical constants) to the thermal Green's function of a network of coupled quantum harmonic oscillators:*

$$G_{pq}(\beta) = e^{-\beta\hbar\Omega}\left(\exp\left[\frac{\beta\hbar\omega^2}{2\Omega}A\right]\right)_{pq}$$

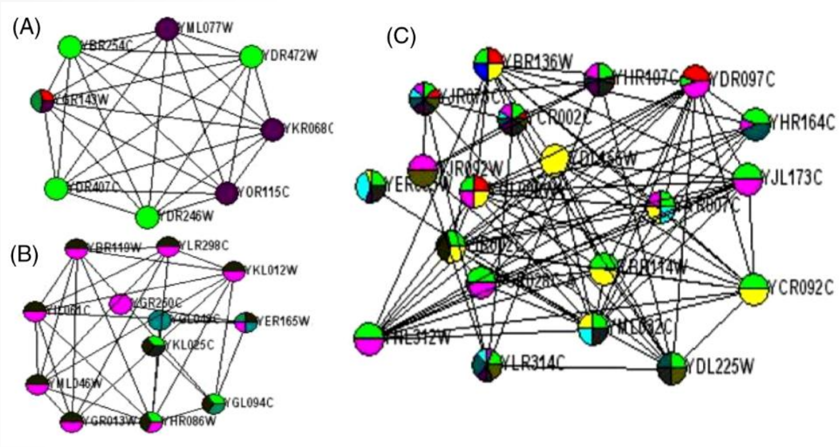
$$\Omega = \sqrt{K/m\omega}$$

$$\beta = \left(k_B T\right)^{-1}$$

Estrada, Hatano & Benzi: *Phys. Rep.* **514** (2012) 89-119.

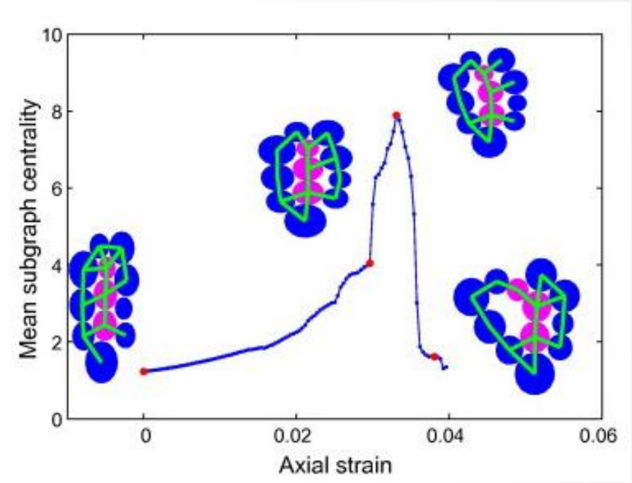
Communicability Applications

Network communities



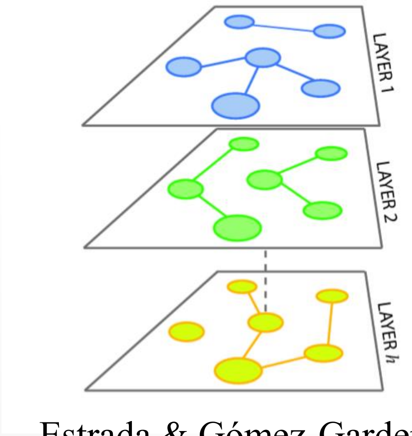
Ma et al: *J. Stat. Mech. Theo. Exp.* (2010) P08012

Granular materials



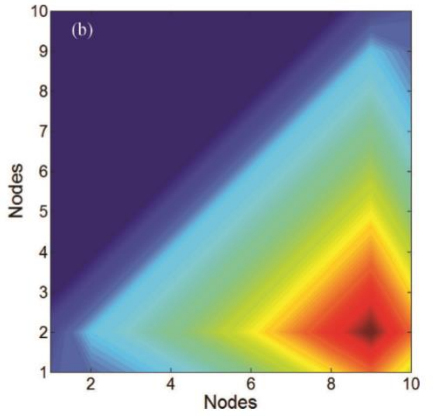
Walker, Tordesillas.:
Int. J. Sol. Struct. 47 (2010) 624-629.

Multiplex networks



Estrada & Gómez-Gardeñes:
Phys. Rev. E 89 (2014) 042819

Temporal networks

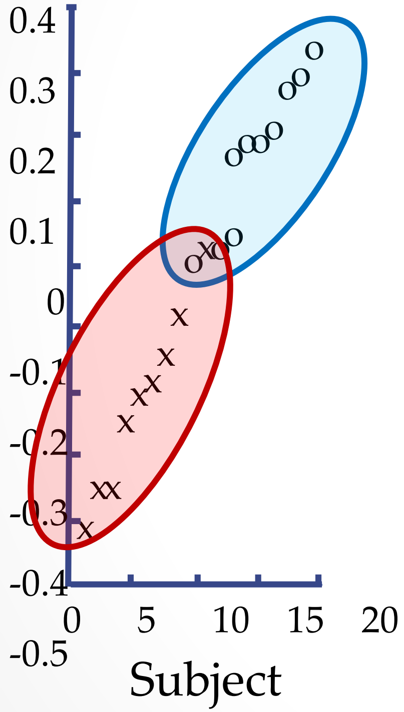


Estrada : *Phys. Rev. E 88* (2013) 042811

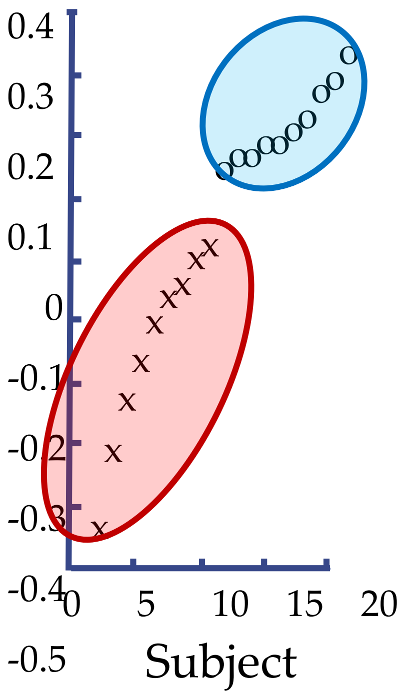
(Neurosciences) Applications

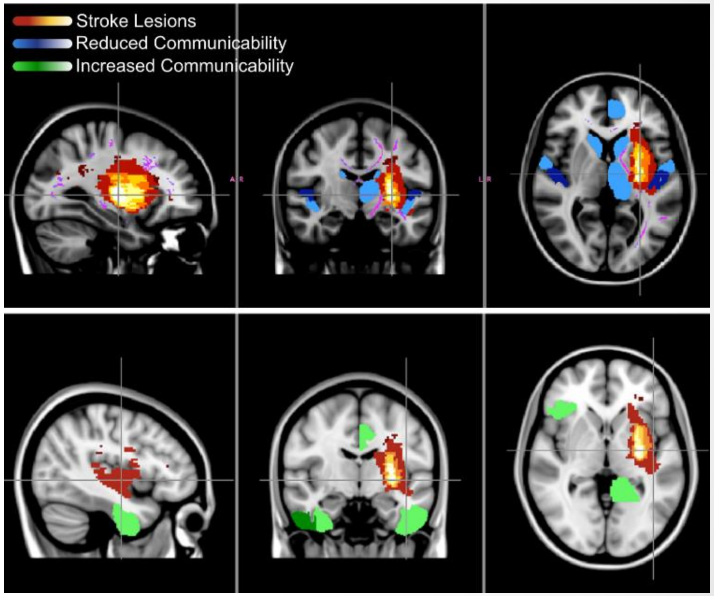
1. Stroke lesions

Shortest path



Communicability





Crofts & Higham: *J. Roy. Soc. Interface* **6** (2009) 411

Crofts et al: *NeuroImage* **54** (2011) 161-9

2. Computational tool

BIOINFORMATICS APPLICATIONS NOTE
Vol. 30 no. 23 2014, pages 3387–3389
doi:10.1093/bioinformatics/btu536

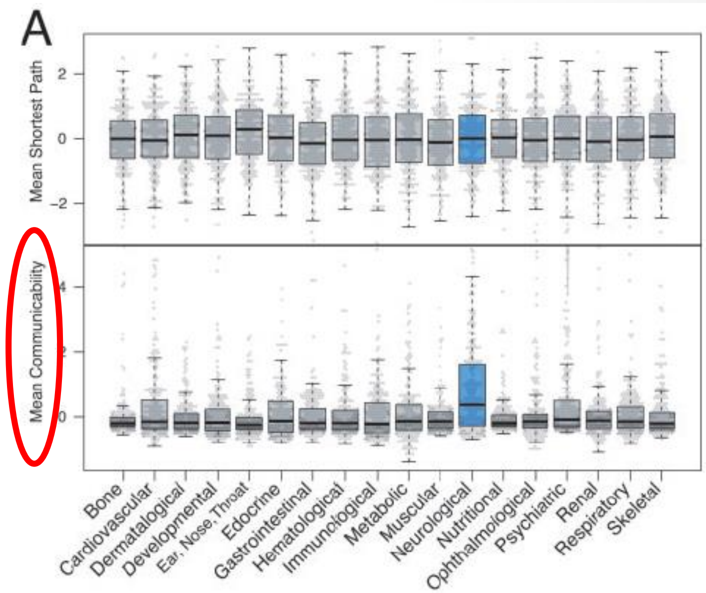
Genome analysis
Advance Access publication August 13, 2014

NetComm: a network analysis tool based on communicability

Ian M. Campbell†, Regis A. James†, Edward S. Chen and Chad A. Shaw*
Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77054, USA
Associate Editor: John Hancock

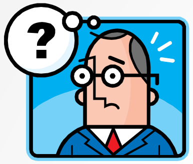
4 CONCLUSIONS

Network communicability provides advantages over alternative metrics because it retains topology information, lends itself to set-based analysis and is easy to represent with univariate scores. It outperforms shortest path in a variety of situations. Overall, our metric may prove useful in the analysis of a variety of biological networks, and our package provides a straightforward approach to computation even on large networks.

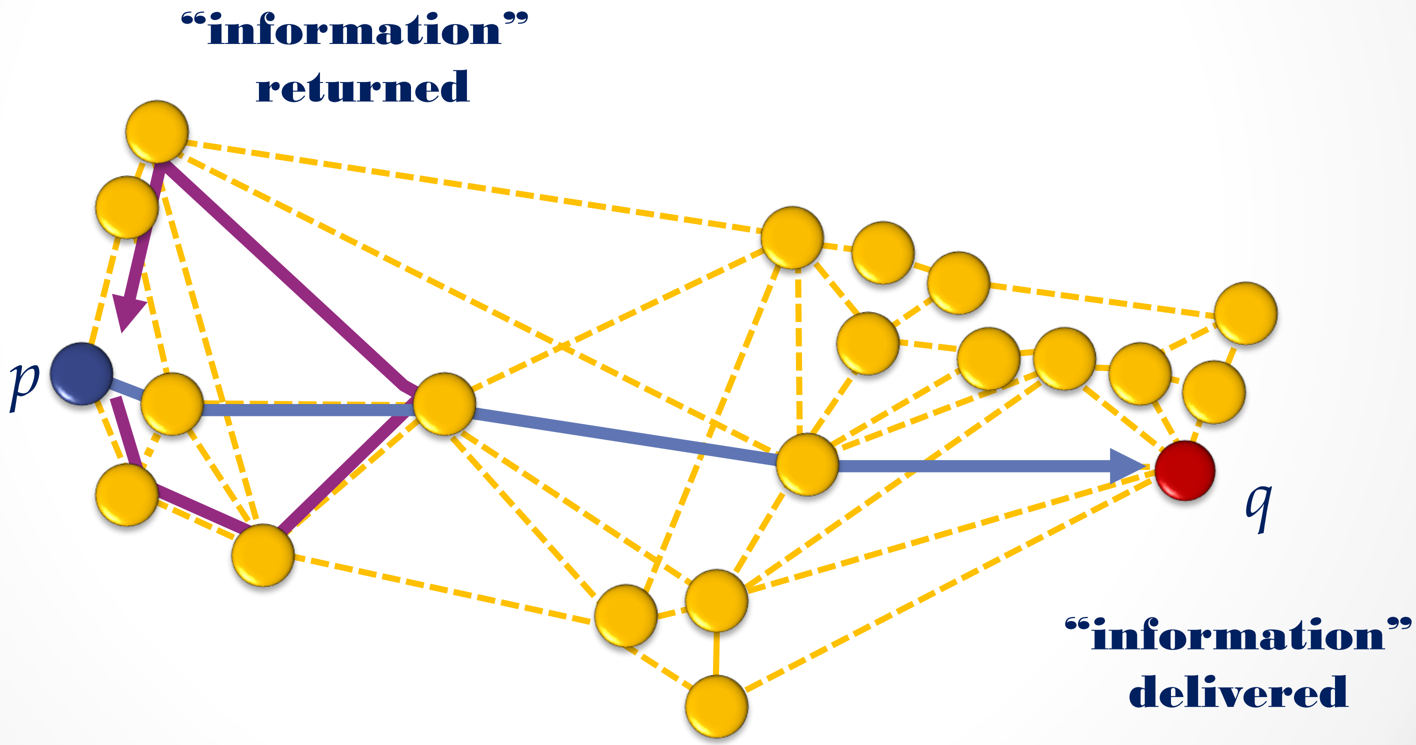


A diversion: Network Bipartivity

Communication Efficiency



In a diffusive-like process, not all the "information" is delivered to the target. Some of it returns to its originator!



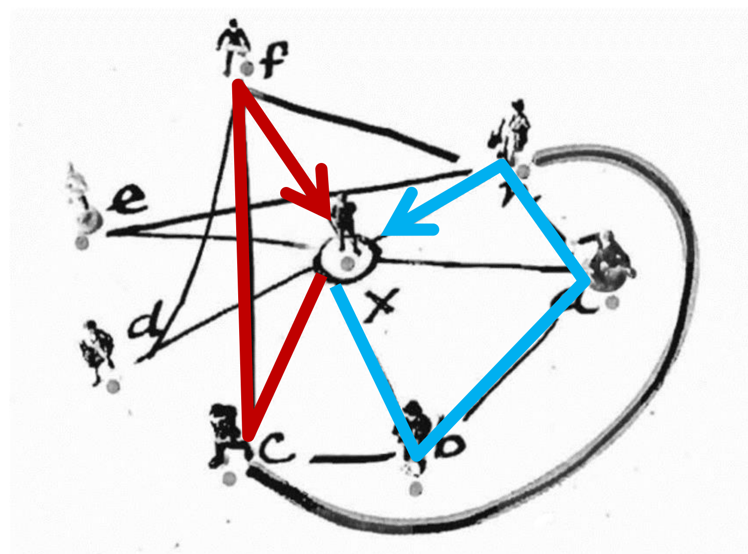
Some evidences…

The Small-World Problem

By Stanley Milgram

Stanley Milgram
1933-1984

With group inbreeding, X's acquaintances feed back into his own circle, normally eliminating new contacts.

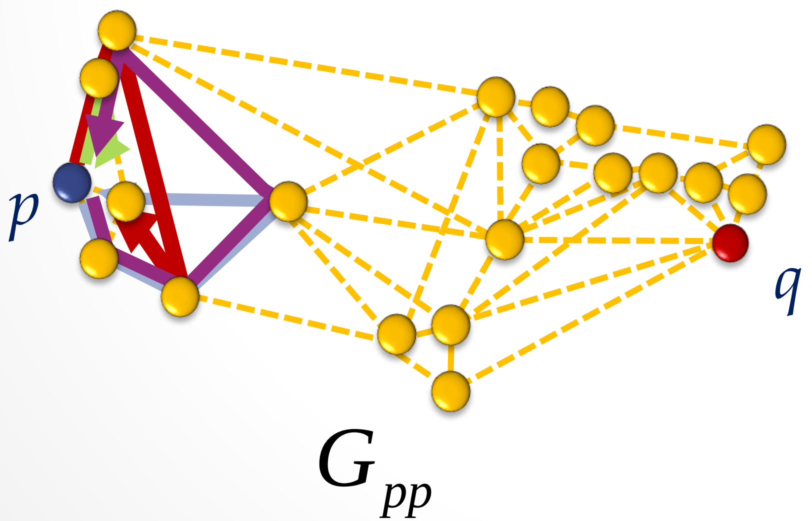


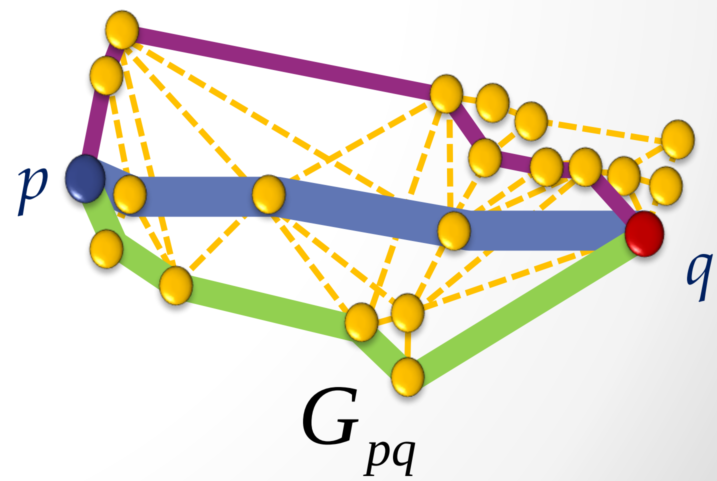
Milgram, S., 1967. *Psychology today*, *2*(1), pp.60-67.

Communication Efficiency

Definition 3. *The "quality of communication" can be quantified as the difference between the number of routes potentially returning the information to each vertex and the number of routes potentially connecting the vertices p and q:*

$$\xi_{pq} = G_{pp} + G_{qq} - 2G_{pq}$$



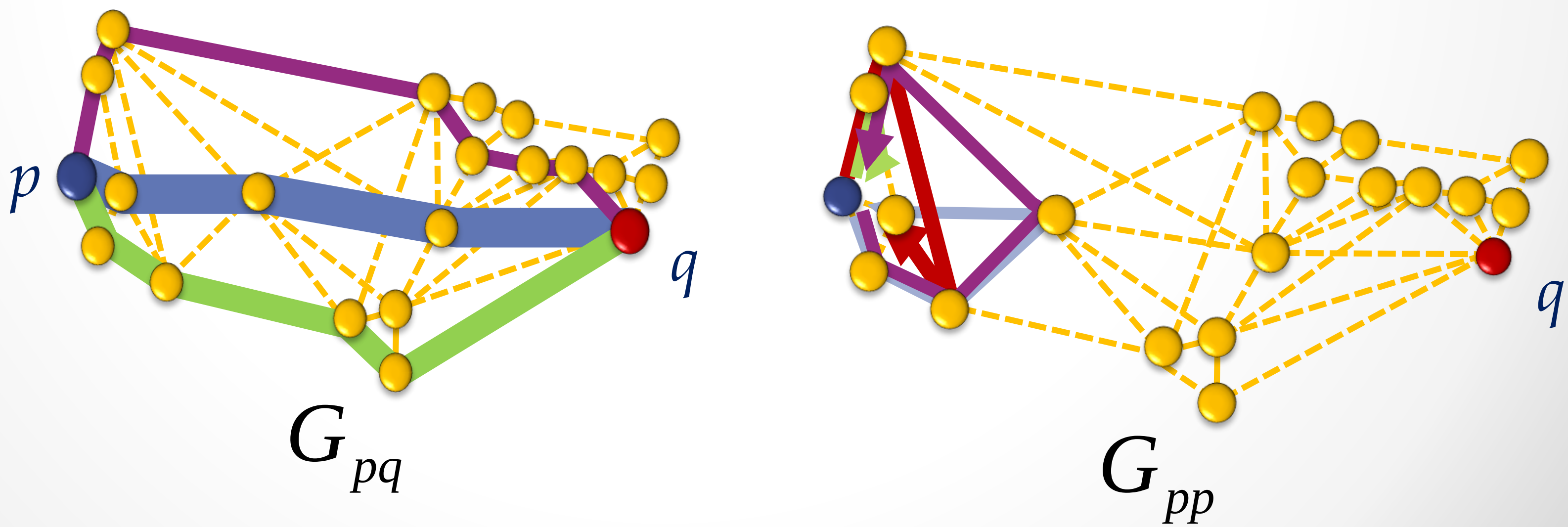


Estrada: *Lin. Alg. Appl*. **436** (2012) 4317-4328.

Communicability Applications

Definition 4. *The "quality of communication" can be quantified as the ratio between the number of routes potentially connecting the vertices p and q, and the number of routes potentially returning the information to each vertex:*

$$\gamma_{pq} = \frac{G_{pq}}{\sqrt{G_{pp}G_{qq}}}$$



Estrada & Hatano: *SIAM Rev*. **58** (2016) 692-715.

Communicability distance

Theorem 2. *The function* ξ_{pq} *is a squared Euclidean distance between the nodes p and q in the network.*

Theorem 3. *The function* γ_{pq} *is the cosine of the Euclidean angle spanned by the position vectors of the nodes p and q.*

$$\theta_{pq} = \cos^{-1} \frac{\vec{x}_p \cdot \vec{x}_q}{\left\|\vec{x}_p\right\| \left\|\vec{x}_q\right\|} = \cos^{-1} \frac{G_{pq}}{\sqrt{G_{pp} G_{qq}}}$$

Estrada: *Lin. Alg. Appl*. **436** (2012) 4317-4328.

Estrada & Hatano: *SIAM Rev*. **58** (2016) 692-715.

Communicability distance

Theorem 4. *The communicability distance induces an embedding of a network into an (n-1)-dimensional Euclidean sphere of radius:*

$$R^2 = \frac{1}{4}\left(c - \frac{(2-b)^2}{a} \right)$$

$$a = \vec{1}^T e^{-A} \vec{1} \qquad b = \vec{s}^T e^{-A} \vec{1} \qquad c = \vec{s}^T e^{-A} \vec{s}$$

Remark. *The communicability distance matrix C is circum-Euclidean:*

$$C = \vec{s}\vec{1}^T + \vec{1}\vec{s}^T - 2e^A \qquad \vec{s} = diag\left(e^A\right)$$

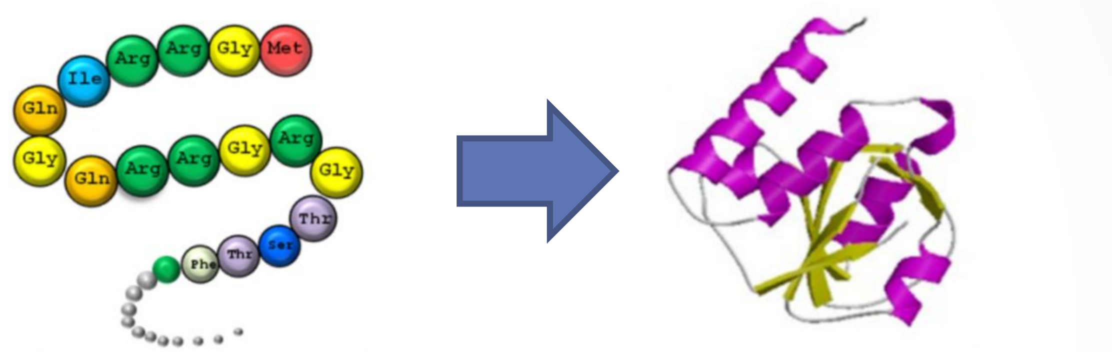
Estrada et al.: *Discrete Appl. Math.* **176** (2014) 53-77.

University of Strathclyde

Induced vs. Imposed Geometry

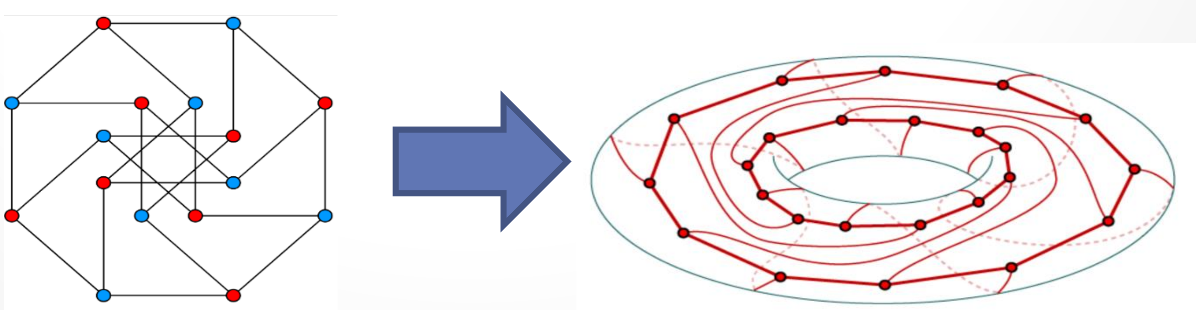
Induced geometry

The geometric properties of the object emerges from the properties of the object itself. The emerged geometry is unique.



Imposed geometry

The geometry appears as a consequence of a given embedding imposed to the object. The geometry is not unique as it depends of the embedding selected.



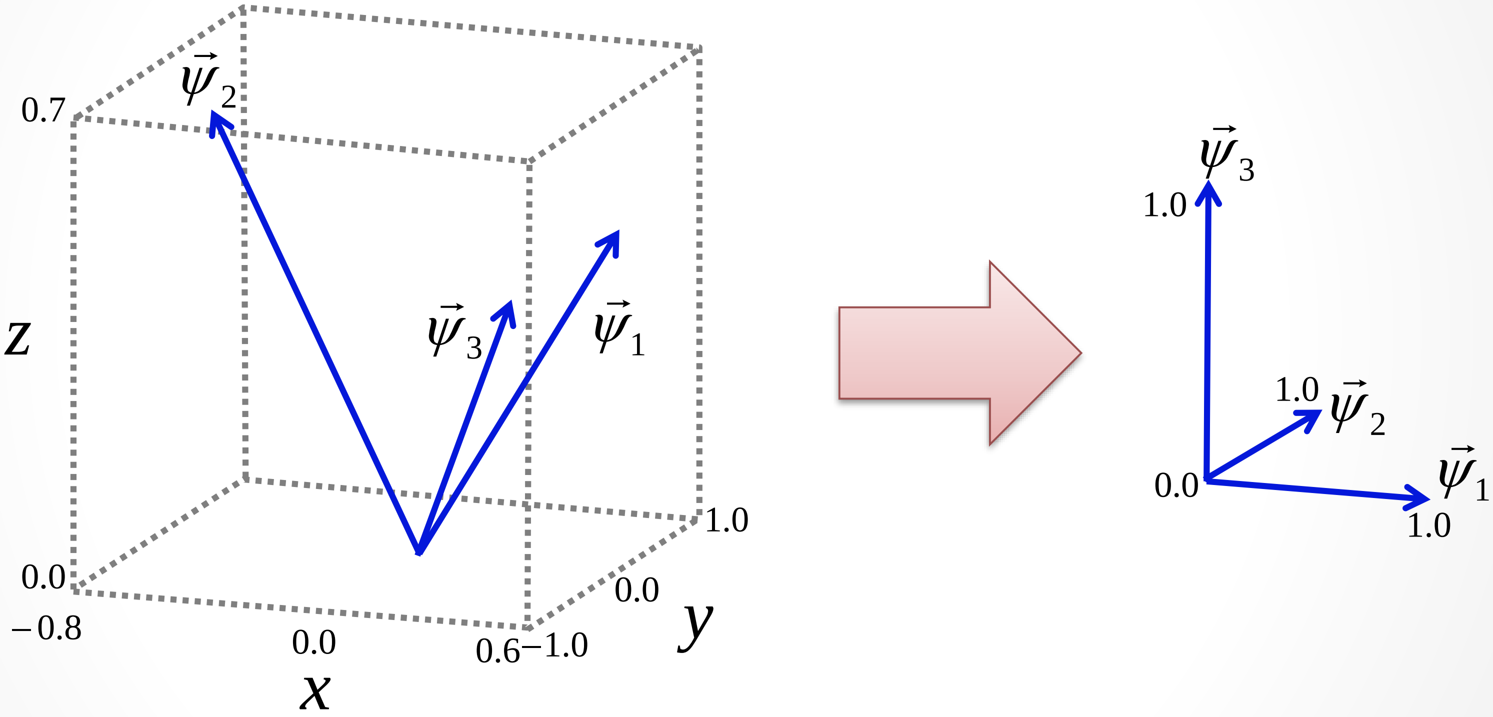
Communicability geometry

Example:

$$A\vec{\psi}_j = \lambda_j \vec{\psi}_j$$

$$\vec{\psi}_1 = \begin{bmatrix} 1/2 \\ 1/\sqrt{2} \\ 1/2 \end{bmatrix} \qquad \vec{\psi}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \qquad \vec{\psi}_3 = \begin{bmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{bmatrix}$$

Communicability Angle



Communicability Angle

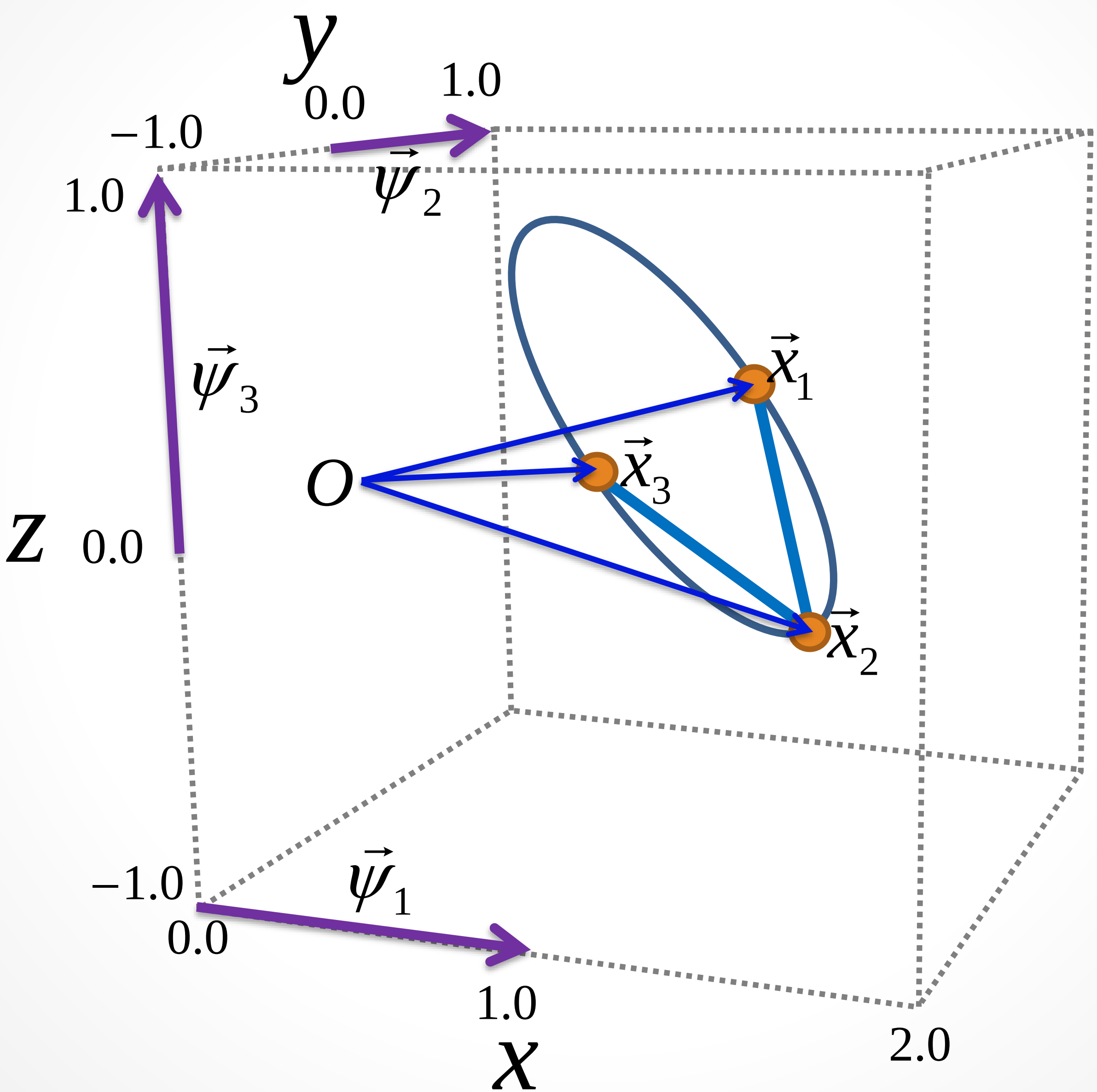
$$A = U \Lambda U^T$$

$$U = \begin{pmatrix} \psi_{1,1} & \psi_{2,1} & \cdots & \psi_{n,1} \\ \psi_{1,2} & \psi_{2,2} & \cdots & \psi_{n,2} \\ \vdots & \vdots & \vdots & \vdots \\ \psi_{1,n} & \psi_{2,n} & \cdots & \psi_{n,n} \end{pmatrix} \qquad \Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

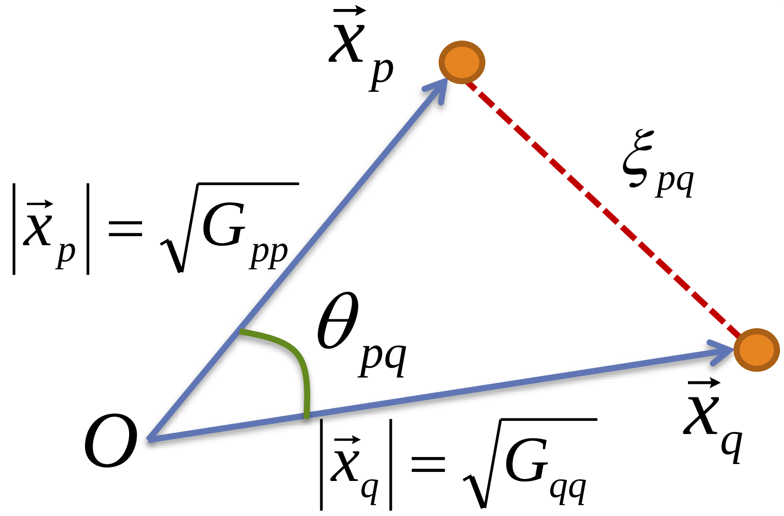
$$\vec{\phi}_p = \left[\psi_{1,p} \cdots \psi_{\mu,p} \cdots \psi_{n,p}\right]^T$$

$$\vec{x}_p = \exp(\Lambda / 2) \vec{\phi}_p$$

Communicability Angle



Communicability Angle



communicability

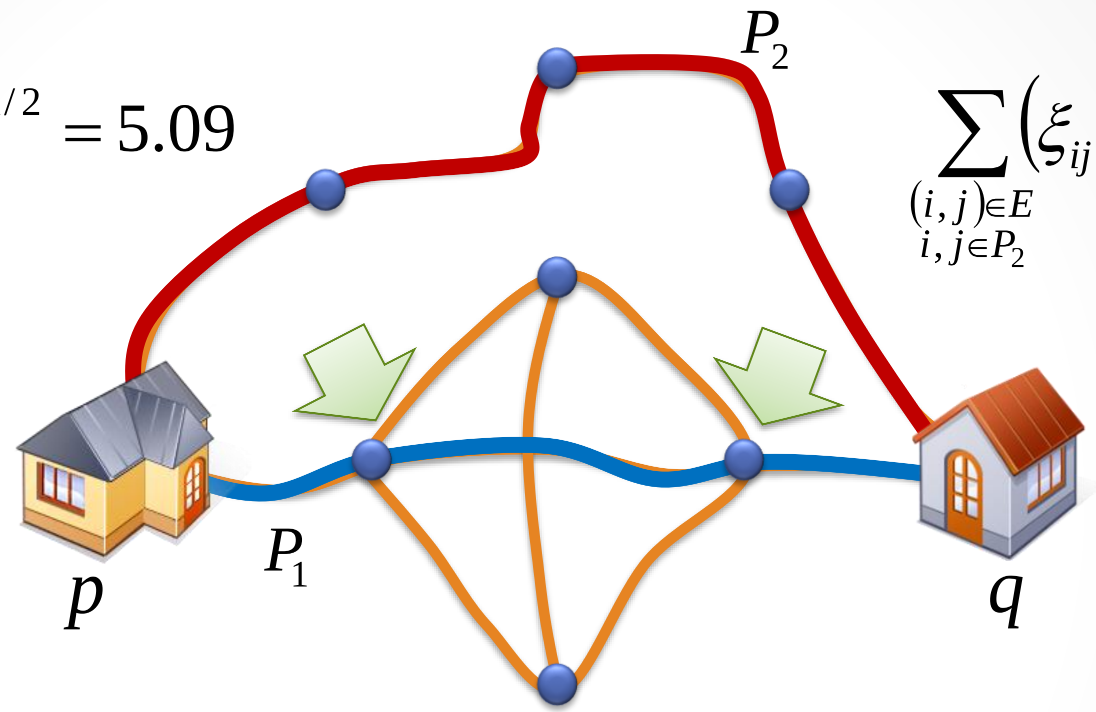
$$G_{pq} = \vec{x}_p \cdot \vec{x}_q$$

Estrada: *Lin. Alg. Appl.* **436** (2012) 4317-4328.
Estrada & Hatano: *SIAM Rev.* **58** (2016) 692-715.

Communicability Distance

$$\sum_{\substack{(i,j)\in E \\ i,j\in P_1}} \left(\xi_{ij}\right)^{1/2} = 5.09$$

$$\sum_{\substack{(i,j)\in E \\ i,j\in P_2}} \left(\xi_{ij}\right)^{1/2} = 4.80$$

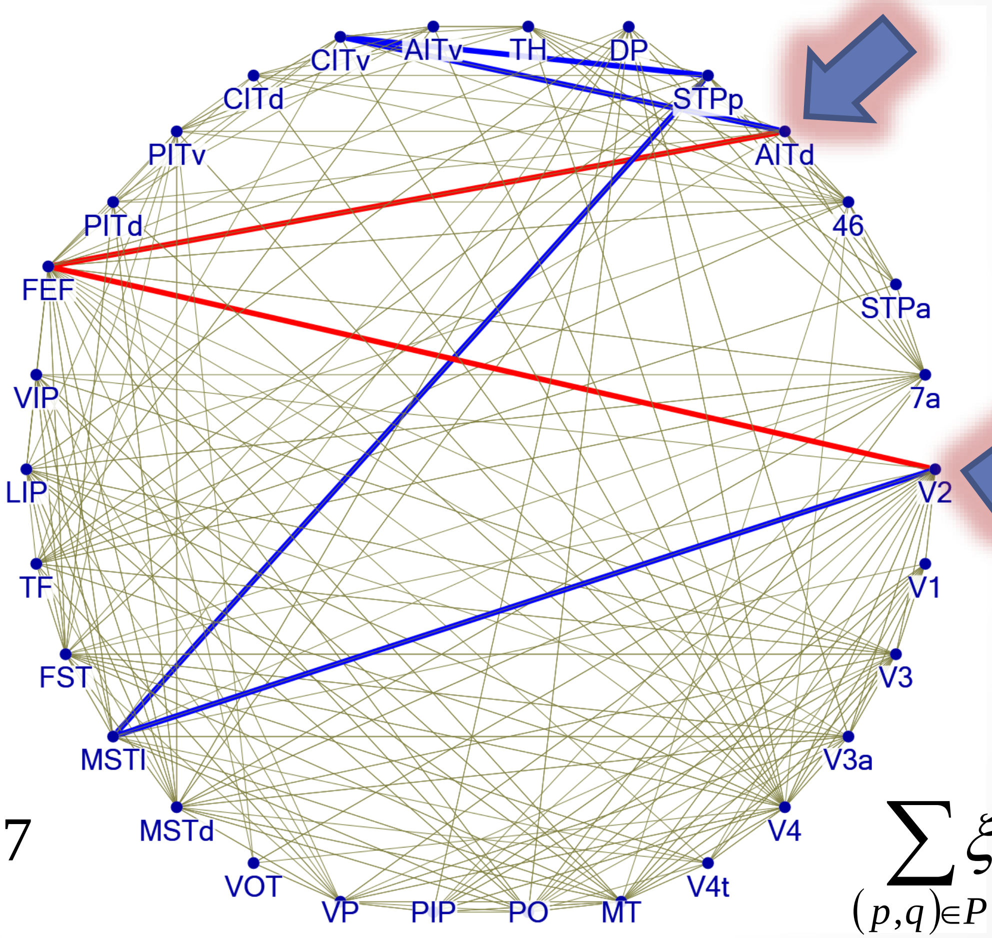


Remark. *The shortest route based on the communicability distance is the shortest path that avoids the nodes with the highest 'returnability' in the graph.*

Estrada: *Phys. Rev. E*. **85** (2012) 066182.

Communicability Distance

macaque visual cortex



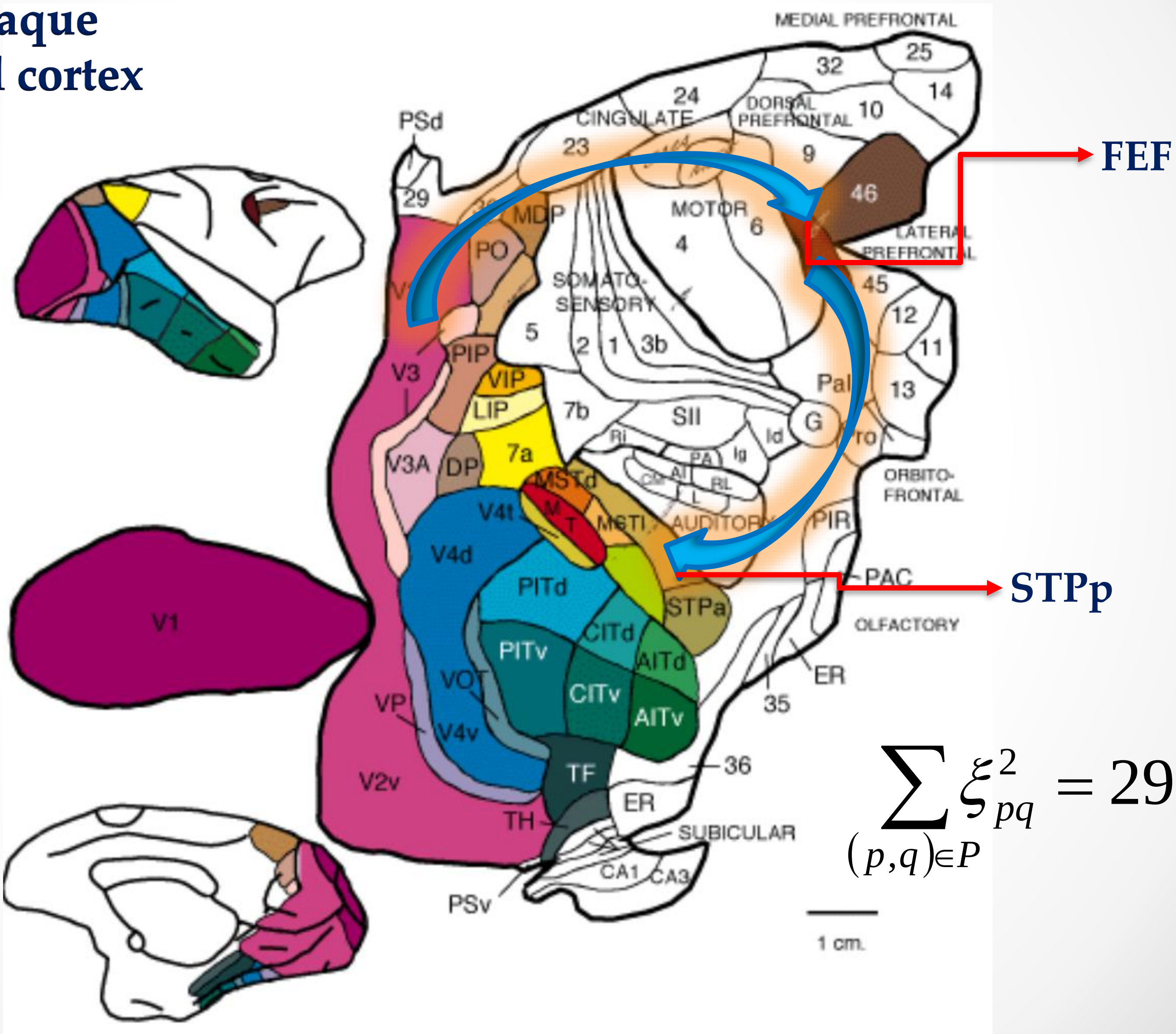
$$\sum_{(p,q)\in P} \xi_{pq}^2 = 295.67$$

$$\sum_{(p,q)\in P} \xi_{pq}^2 = 121.86$$

Data: Sporn & Kotter. *Plos Biol., 2* (2004), e369.

Communicability Distance

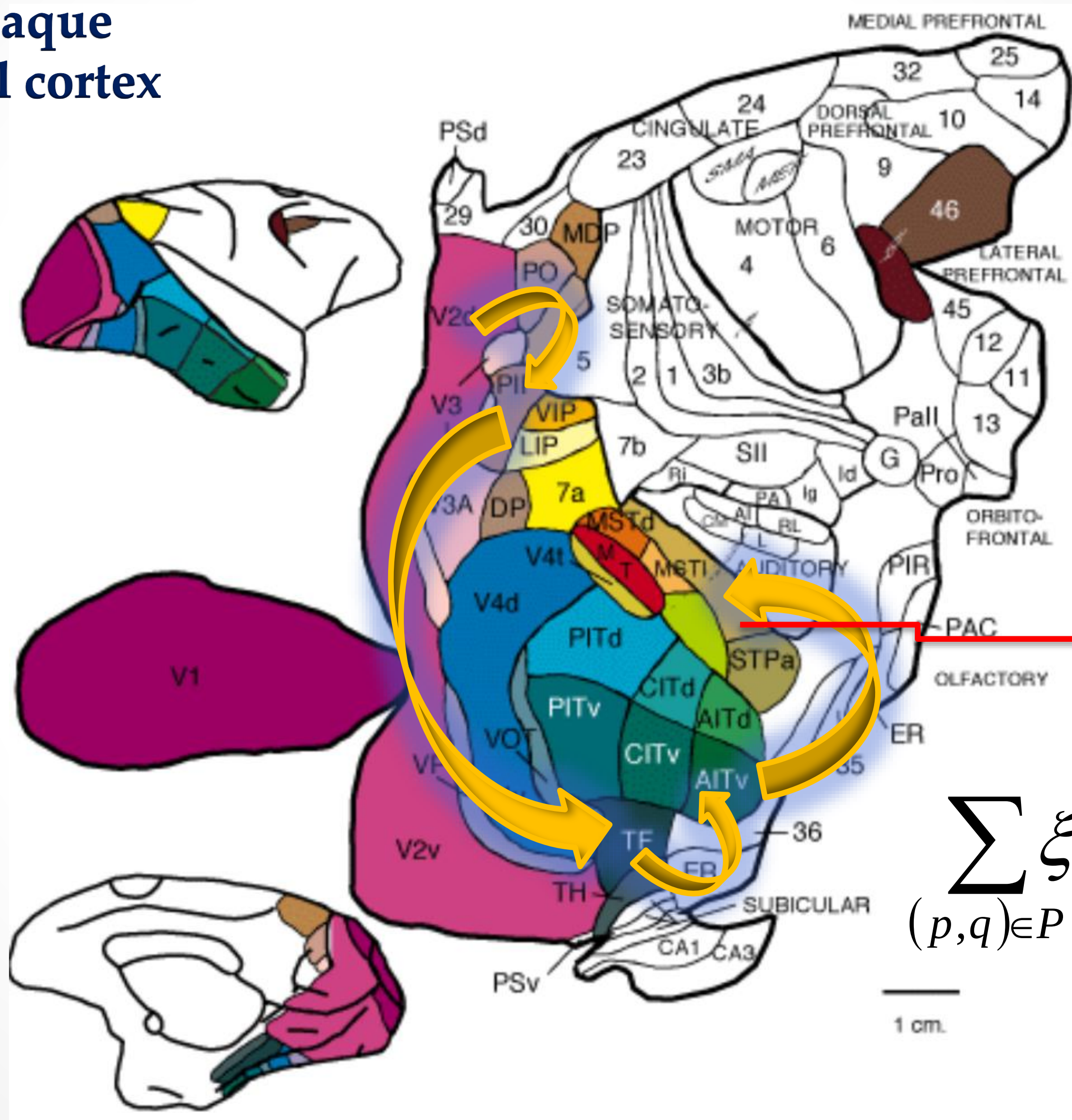
macaque visual cortex



$$\sum_{(p,q)\in P} \xi_{pq}^{2} = 295.67$$

Communicability Distance

macaque visual cortex

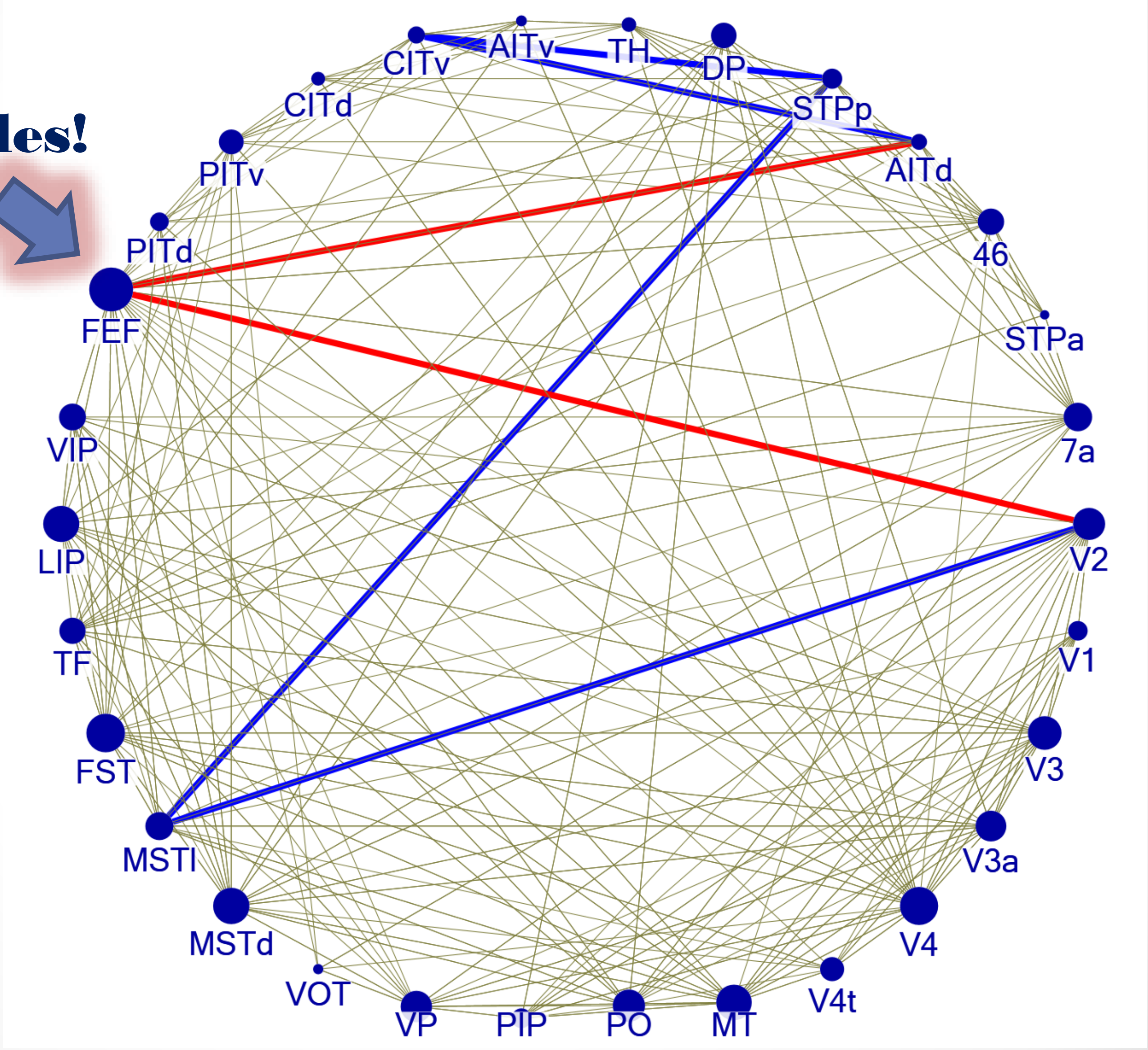


STPp

$$\sum_{(p,q)\in P} \xi_{pq}^{2} = 121.86$$

Communicability Distance

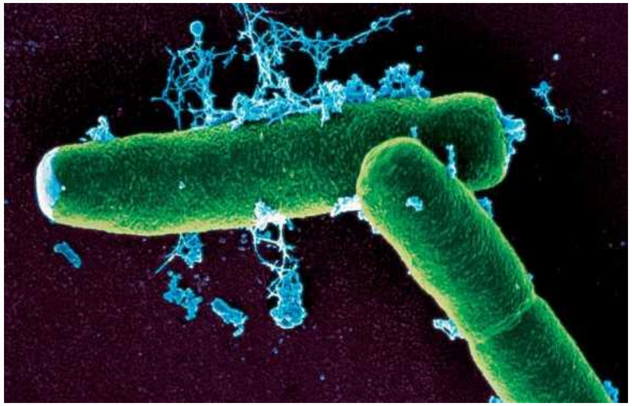
113 triangles!



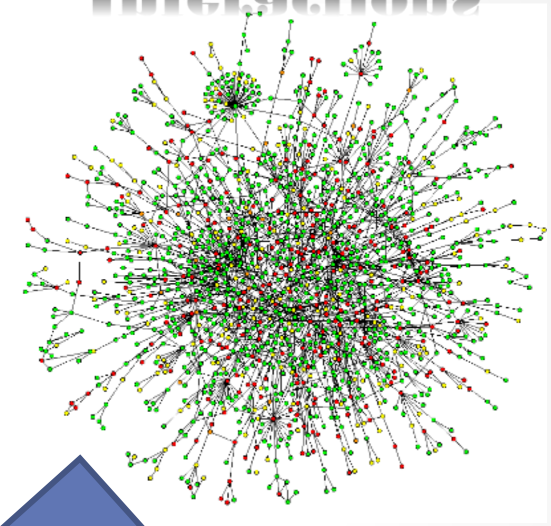
Data: Sporn & Kotter. *Plos Biol., 2* (2004), e369.

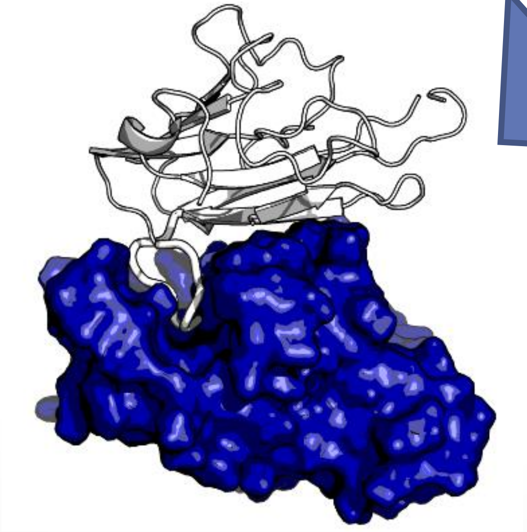
Closeness centralities

Yeast



Protein-protein Interactions



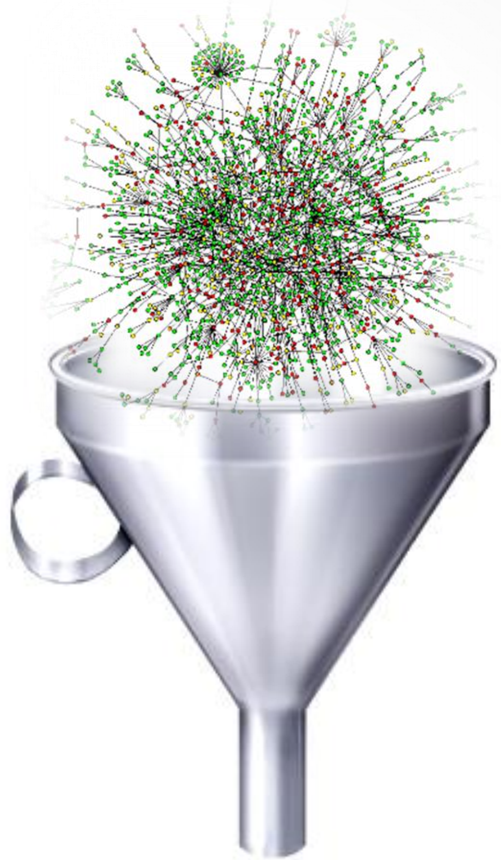


Essential

Closeness centralities

PPI network

Centrality measure

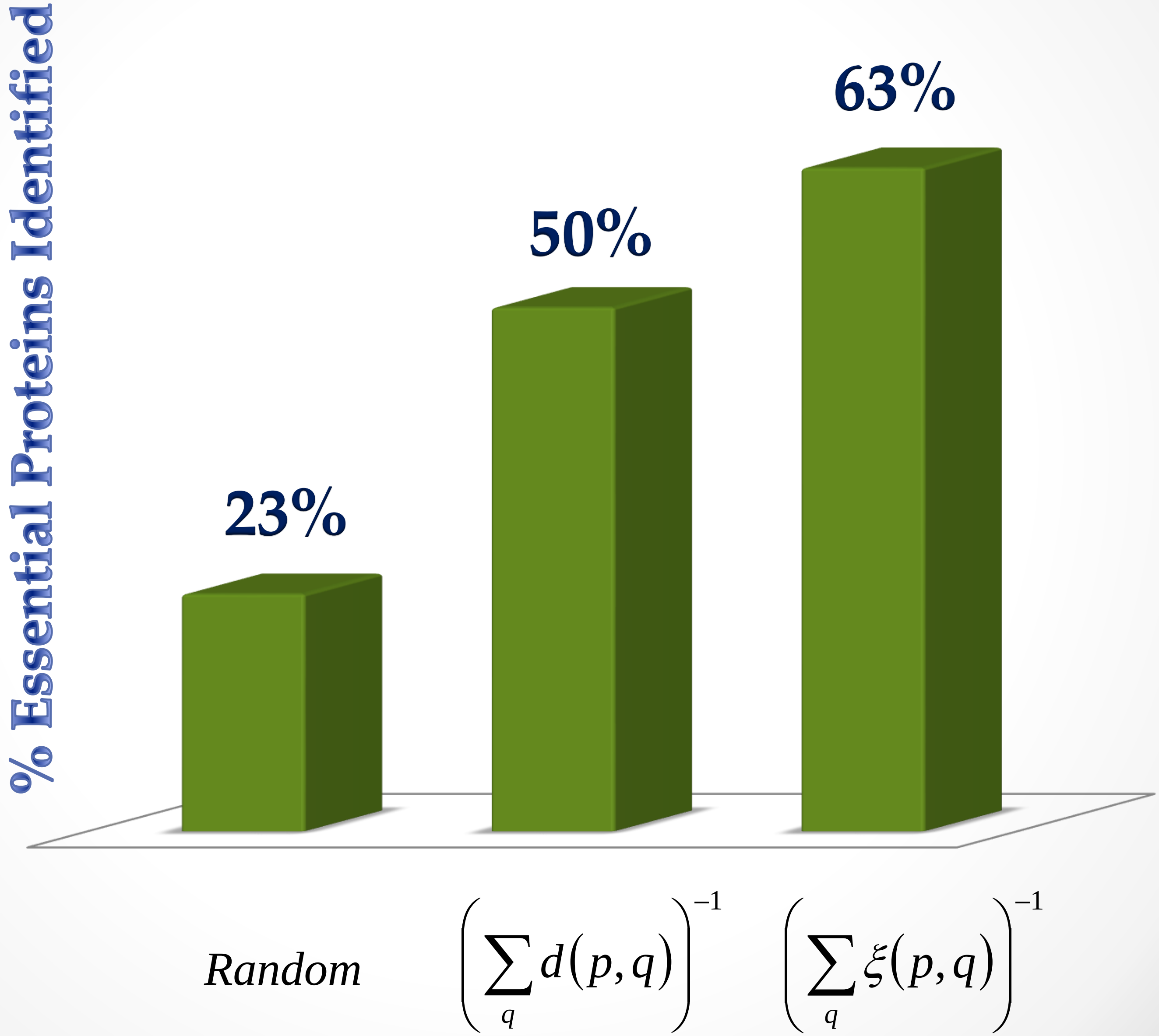


Ranking

Rank	Protein	Essential ?
1	YCR035C	Y
2	YIL062C	Y
...	...	...

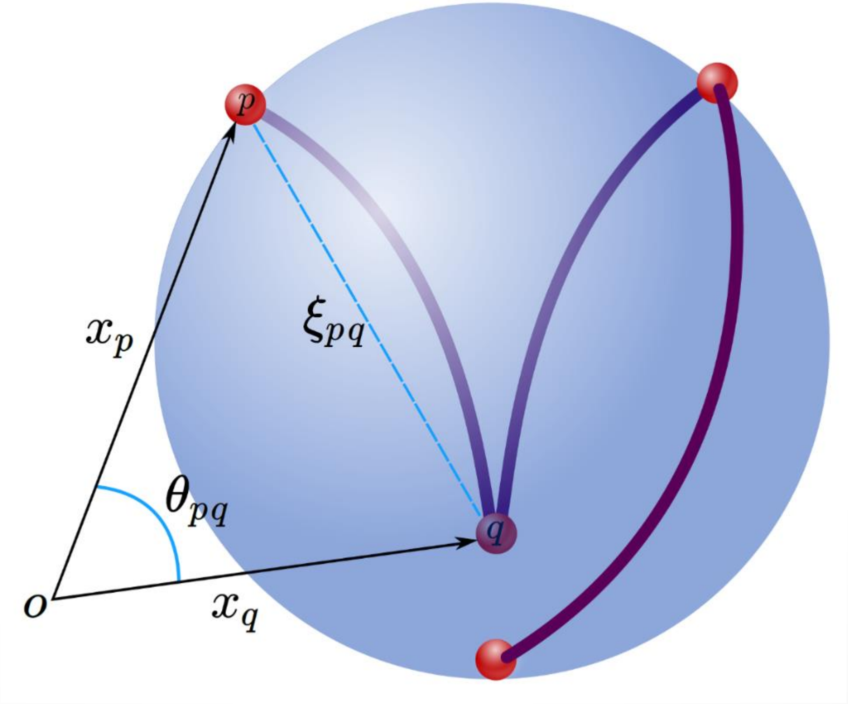
Estrada: *Proteomics* **6** (2006) 35-40.

Closeness centralities



Network 'Spatial Efficiency'

Definition 4. *Let $\chi(G)$ be a metric space defined by the 'flow of communication' among the vertices in the network G. The spatial efficiency is measured by the packing of the vertices of G in $\chi(G)$.*

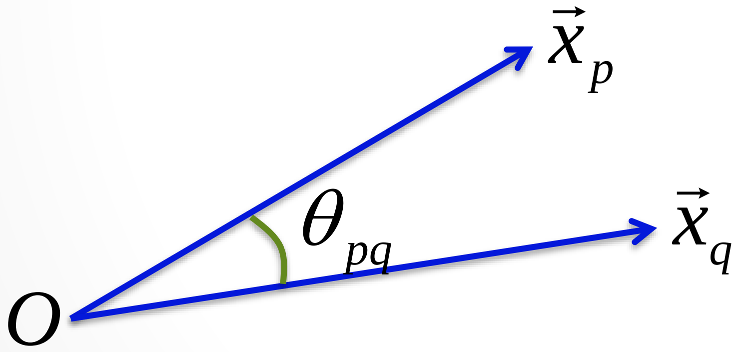


$\langle\theta\rangle$ and Spatial Efficiency

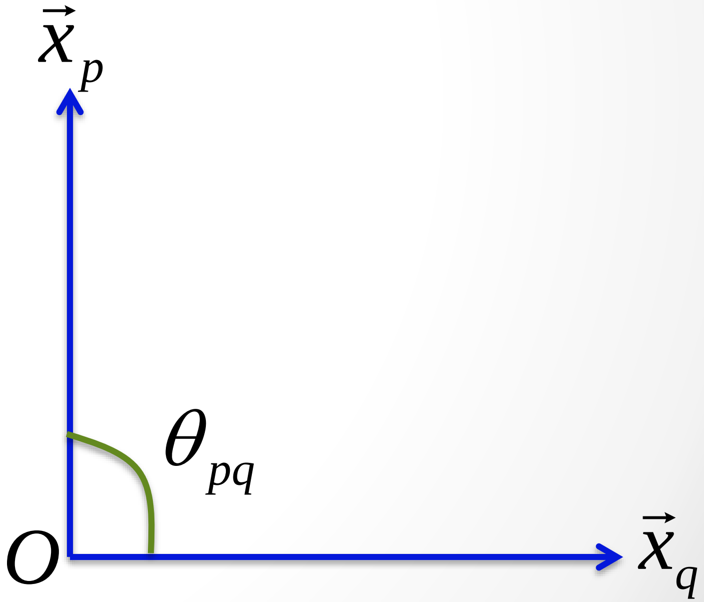
$\langle\theta\rangle$ *: average communicability angle*

For simple, unweighted, undirected networks:

$$0^\circ \le \theta_{pq} \le 90^\circ$$

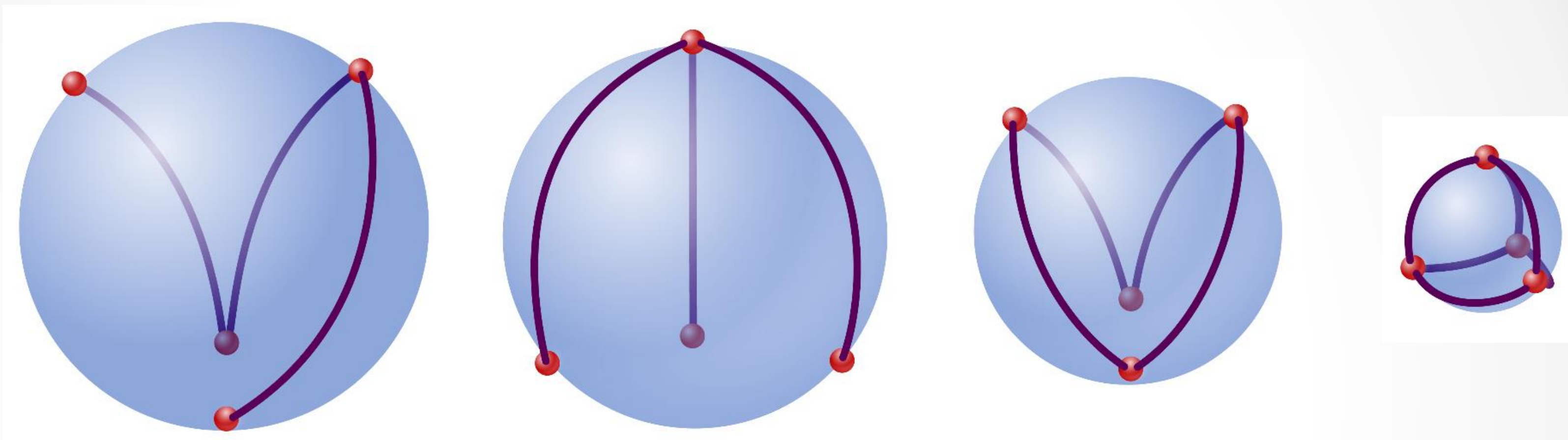


Good spatial efficiency



Poor spatial efficiency

Network 'Spatial Efficiency'



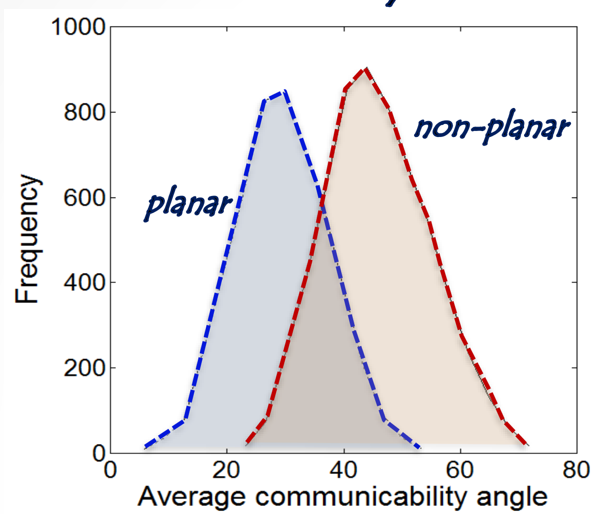
$\langle\theta\rangle = 0.867$ $\langle\theta\rangle = 0.817$ $\langle\theta\rangle = 0.731$ $\langle\theta\rangle = 0.525$

Spatial efficiency

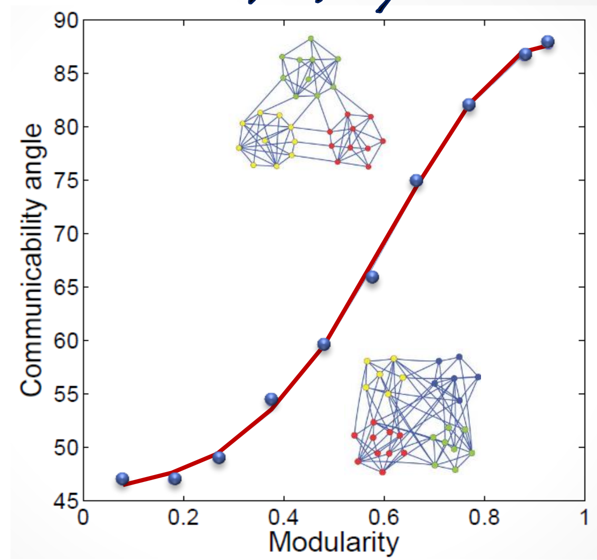
Estrada & Hatano: *SIAM Rev*. **58** (2016) 692-715.

Network 'Spatial Efficiency'

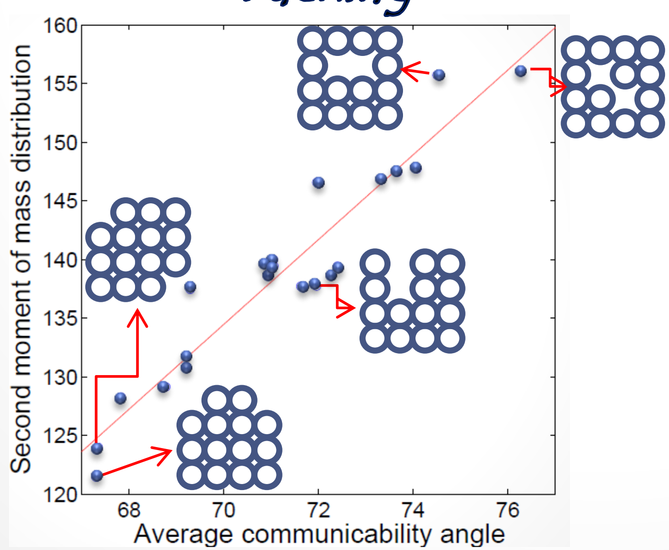
Planarity



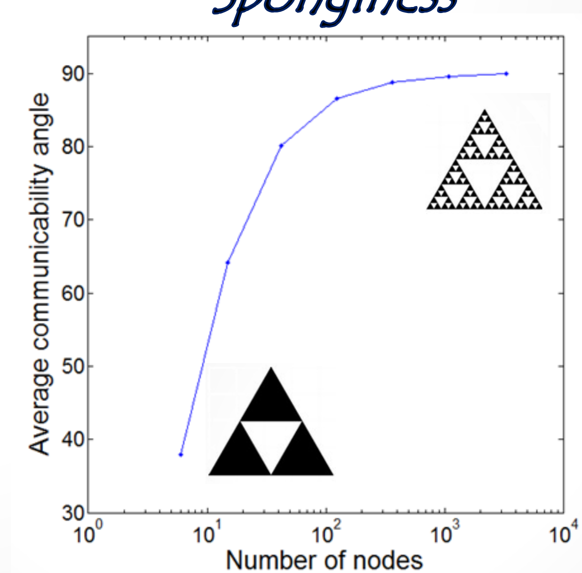
Modularity



Packing

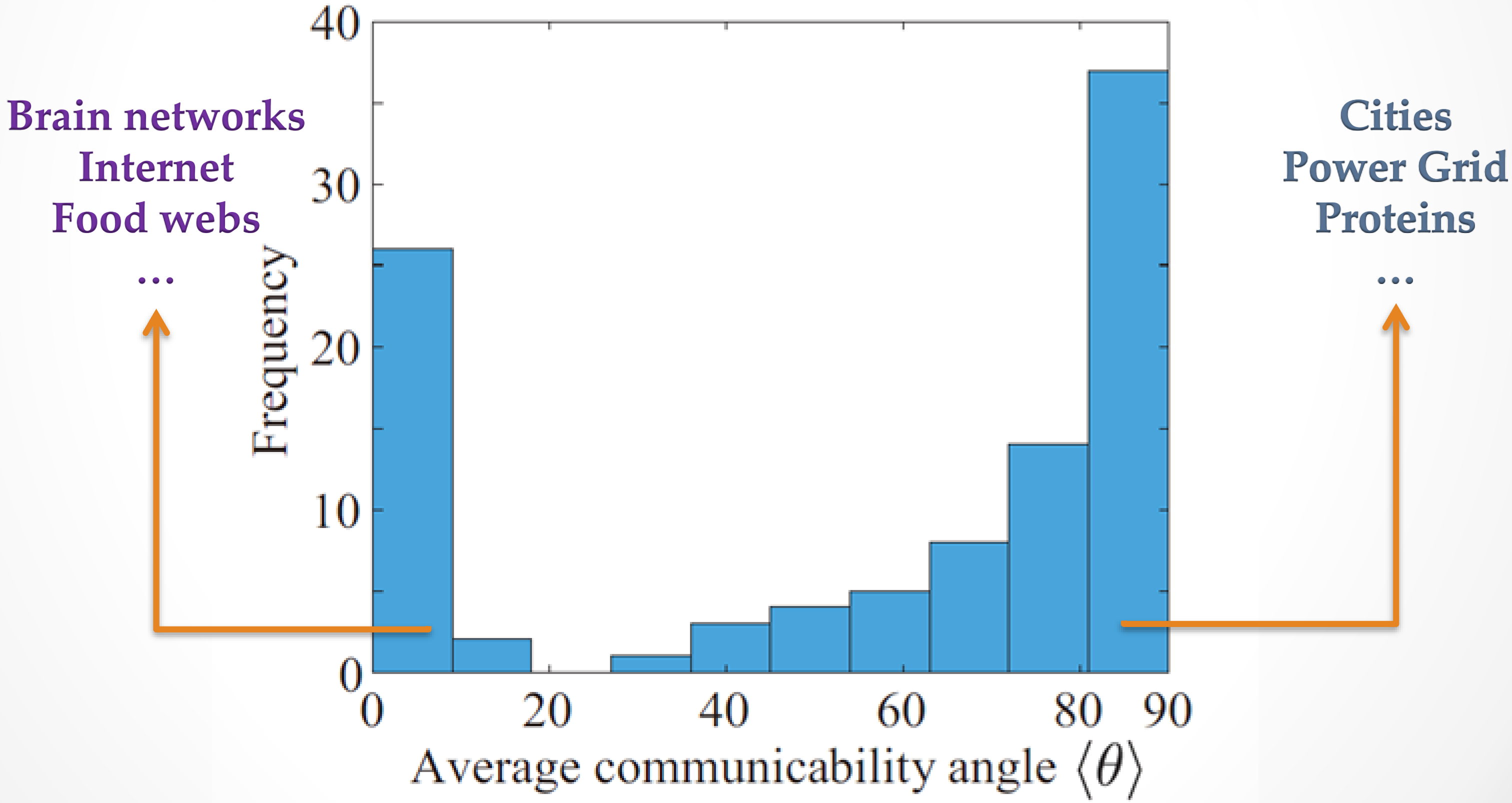


'Sponginess'



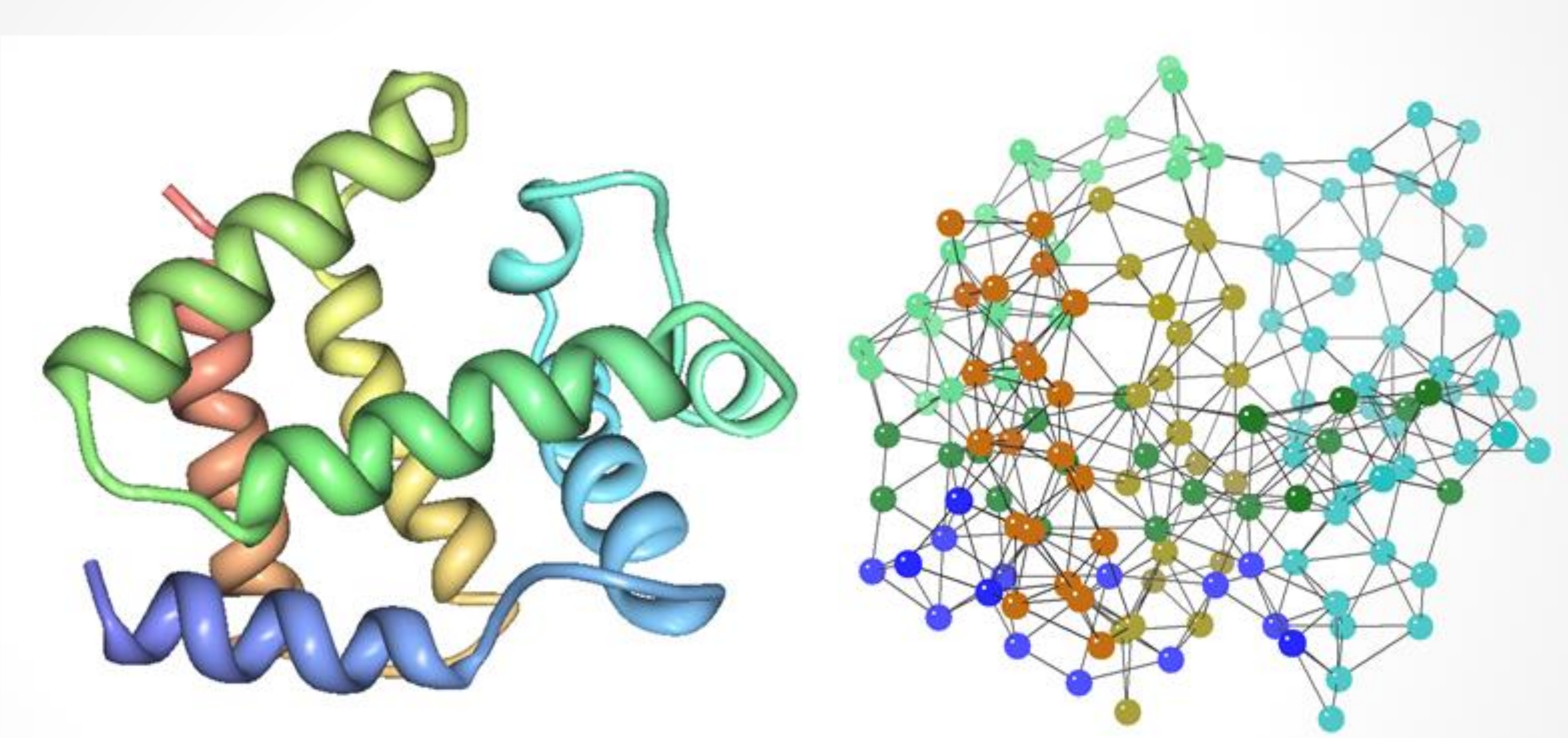
Spatial Efficiency in the Real-World

We studied 120 real-world networks representing biological, ecological, infrastructural, social, and technological systems.



Estrada & Hatano: *SIAM Rev*. **58** (2016) 692-715.

$\langle \theta \rangle$ in the Real-World



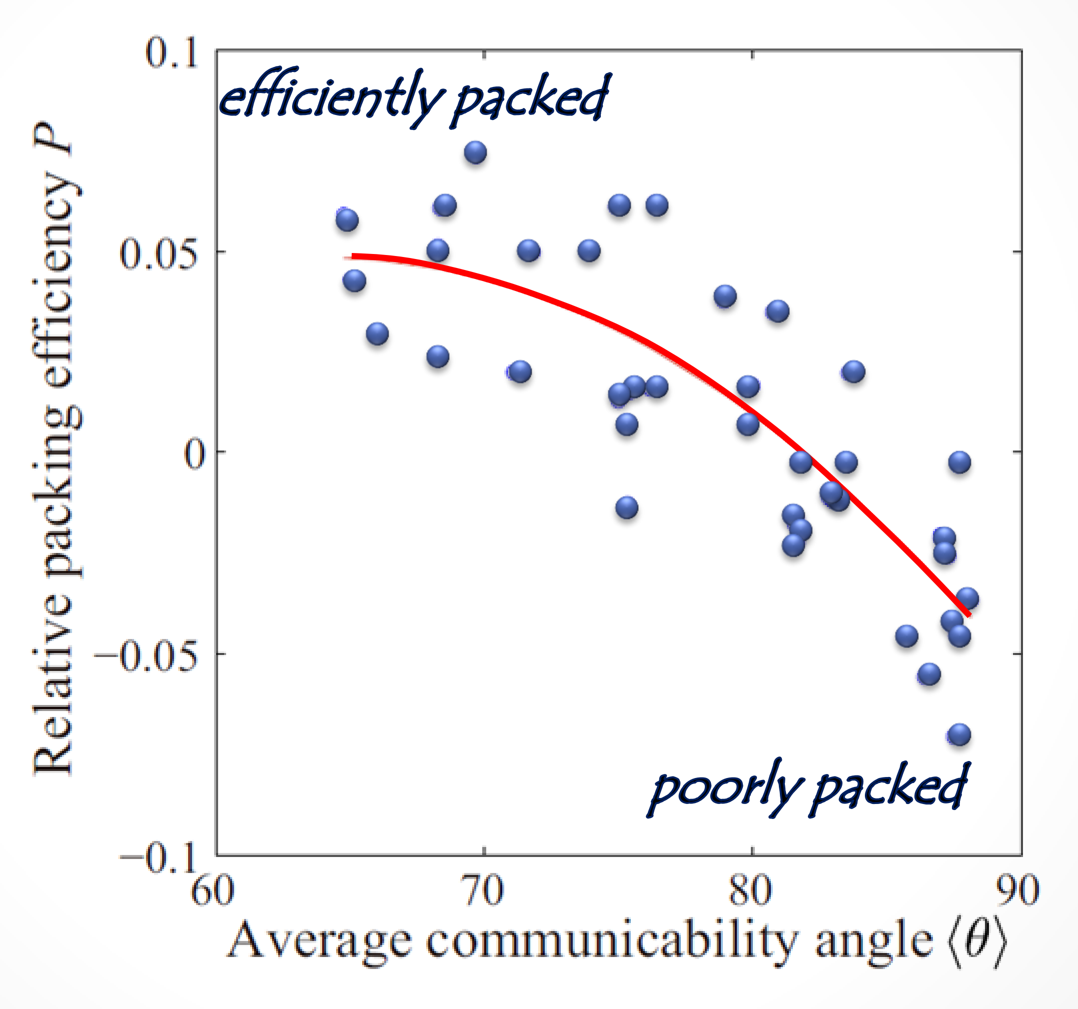
Relative Packing Efficiency

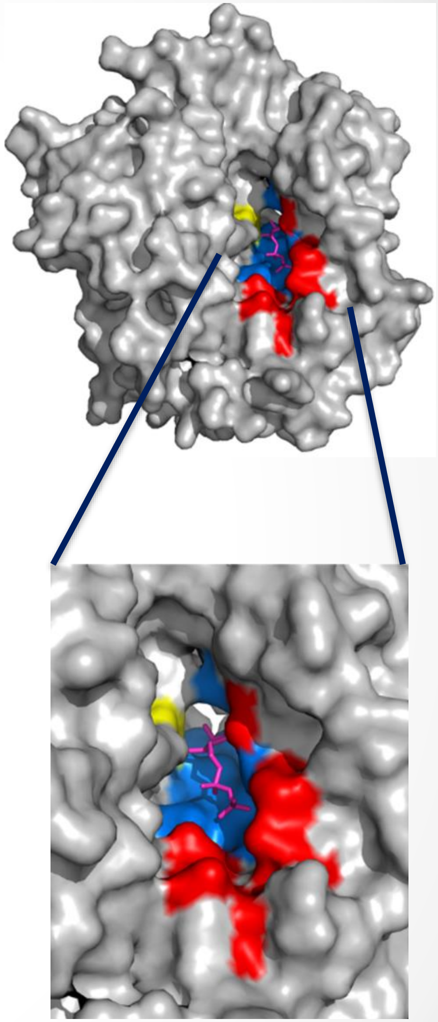
$$P = \frac{V_e - V_o}{V_e}$$

V_e expected volume from ideal 3D structure of the protein

V_o observed volume from X-rays crystal of the protein

$\langle\theta\rangle$ in the Real-World



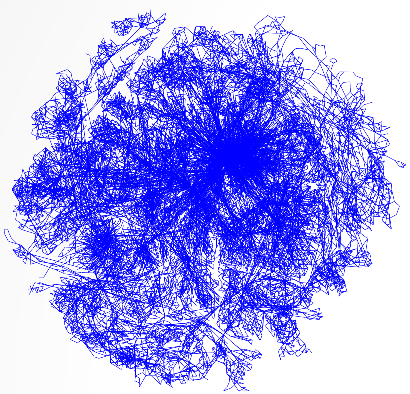


Estrada & Hatano: *SIAM Rev*. **58** (2016) 692-715.

Spatial Efficiency in the Real-World

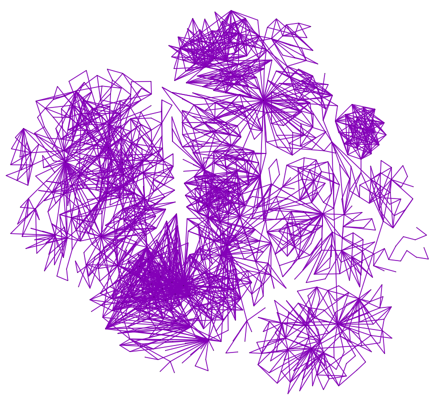
1) Spatially embedded networks

Barcelona



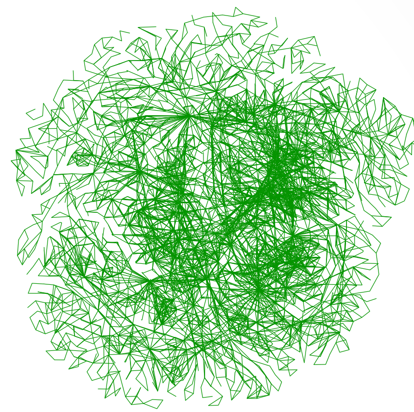
$\langle\theta\rangle = 71.9^{\circ}$

Rio Grande



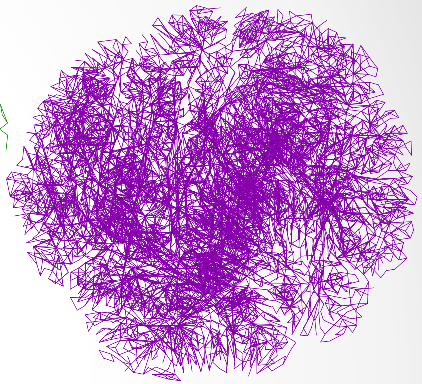
$\langle\theta\rangle = 79.7^{\circ}$

Atlanta



$\langle\theta\rangle = 86.5^{\circ}$

Berlin

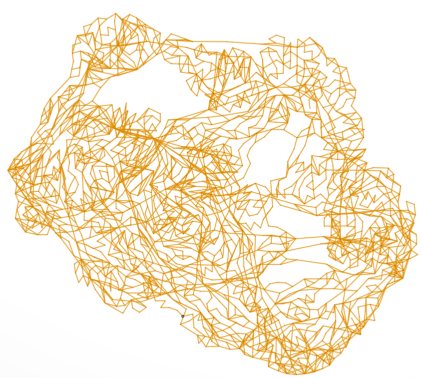


$\langle\theta\rangle = 88.2^{\circ}$

Hong Kong

$\langle\theta\rangle = 88.9^{\circ}$

Mecca

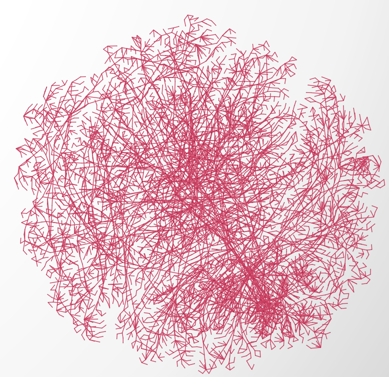


$\langle\theta\rangle = 89.5^{\circ}$

Oxford

$\langle\theta\rangle = 89.5^{\circ}$

Milton Keynes



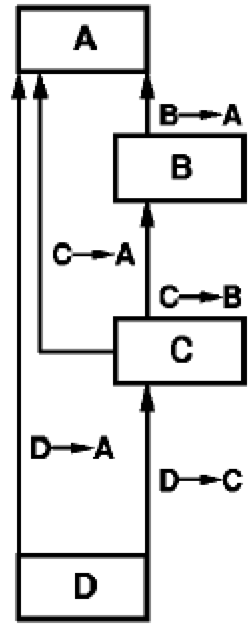
$\langle\theta\rangle = 89.9^{\circ}$

Spatial Efficiency in the Real-World

2) Non geographically embedded networks

Class Collaboration Graph

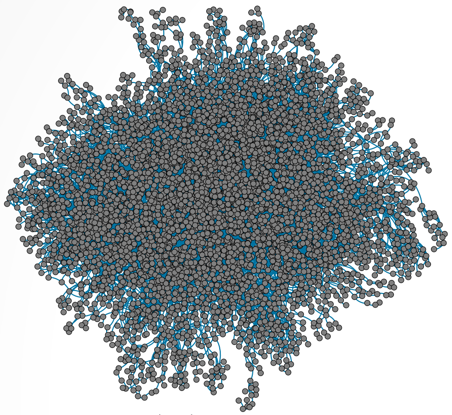
```
class A {
        // definition of class A
};
class B {
        A* ab;
        // rest of definition of class B
};
class C {
        A* ac;
        B* bc;
        // rest of definition of class C
};
class D: public C {
        A* ad;
        // rest of definition of class D
};
```



Myers: *Phys. Rev. E* **68** (2003) 046116.

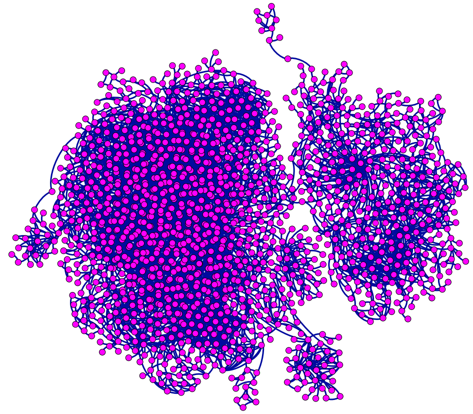
Spatial Efficiency in the Real-World

Linux



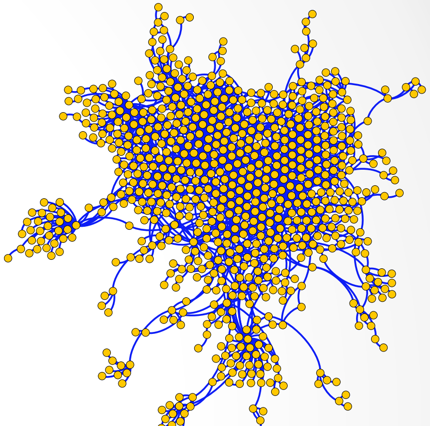
$\langle \theta \rangle = 3.5^{\circ}$

MySQL



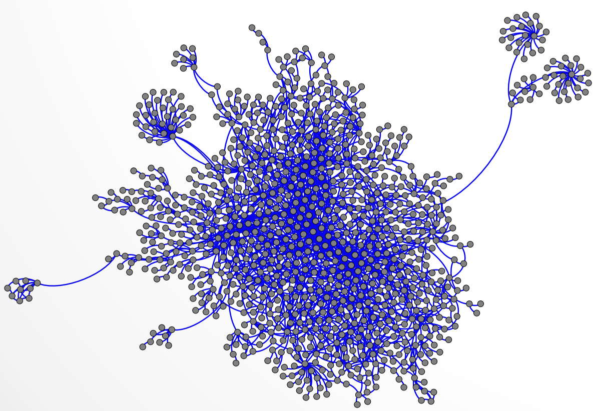
$\langle \theta \rangle = 45.7^{\circ}$

VTK



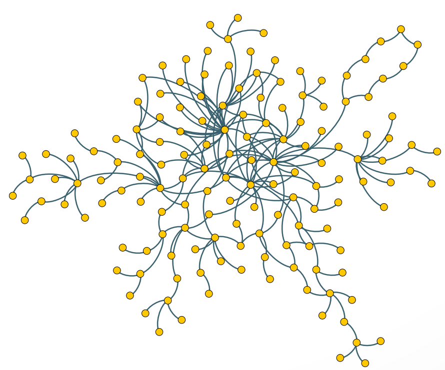
$\langle \theta \rangle = 70.1^{\circ}$

AbiWord



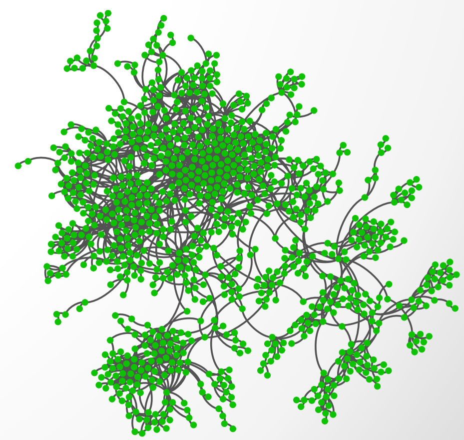
$\langle \theta \rangle = 72.9^{\circ}$

Digital Material



$\langle \theta \rangle = 81.6^{\circ}$

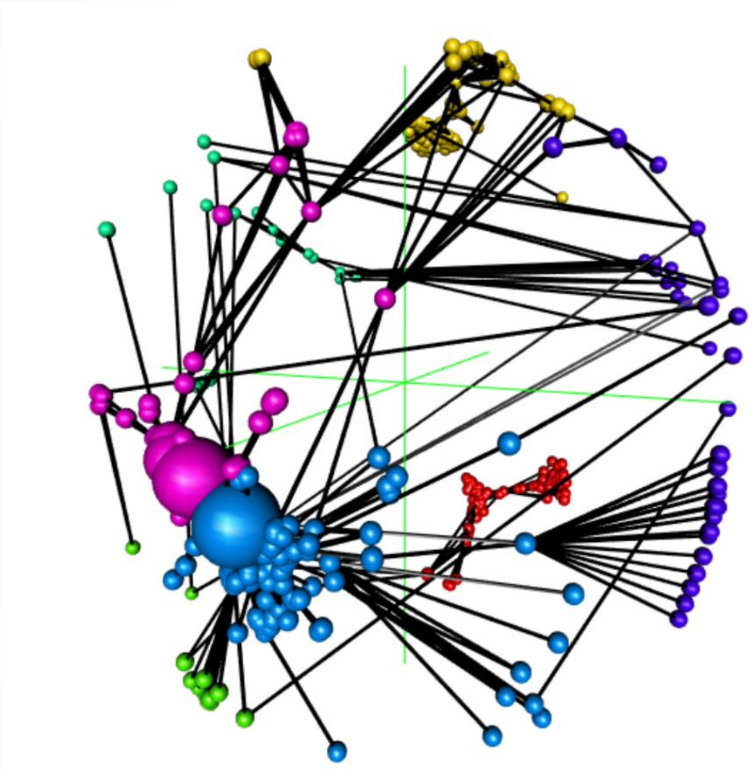
XMMS

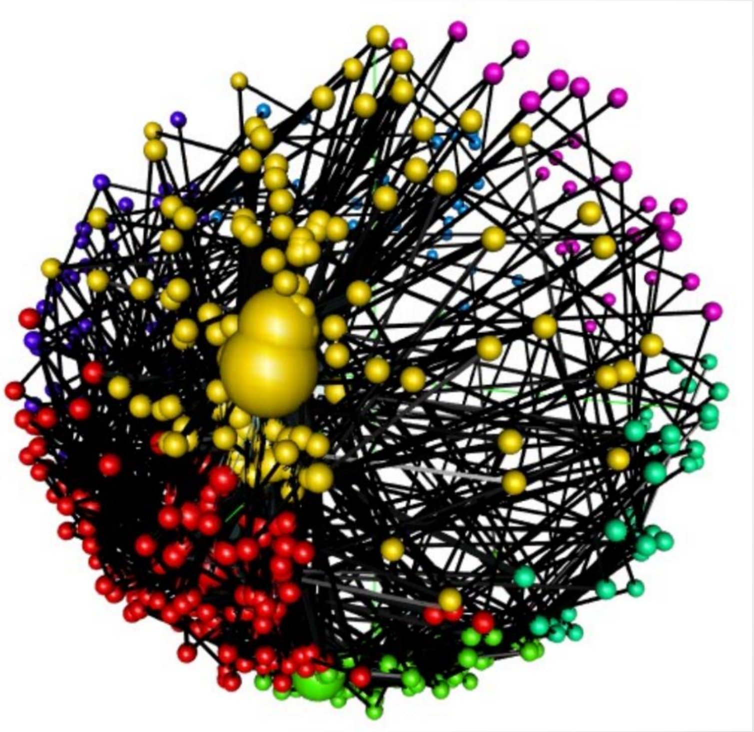


$\langle \theta \rangle = 84.3^{\circ}$

Reducing dimensionality

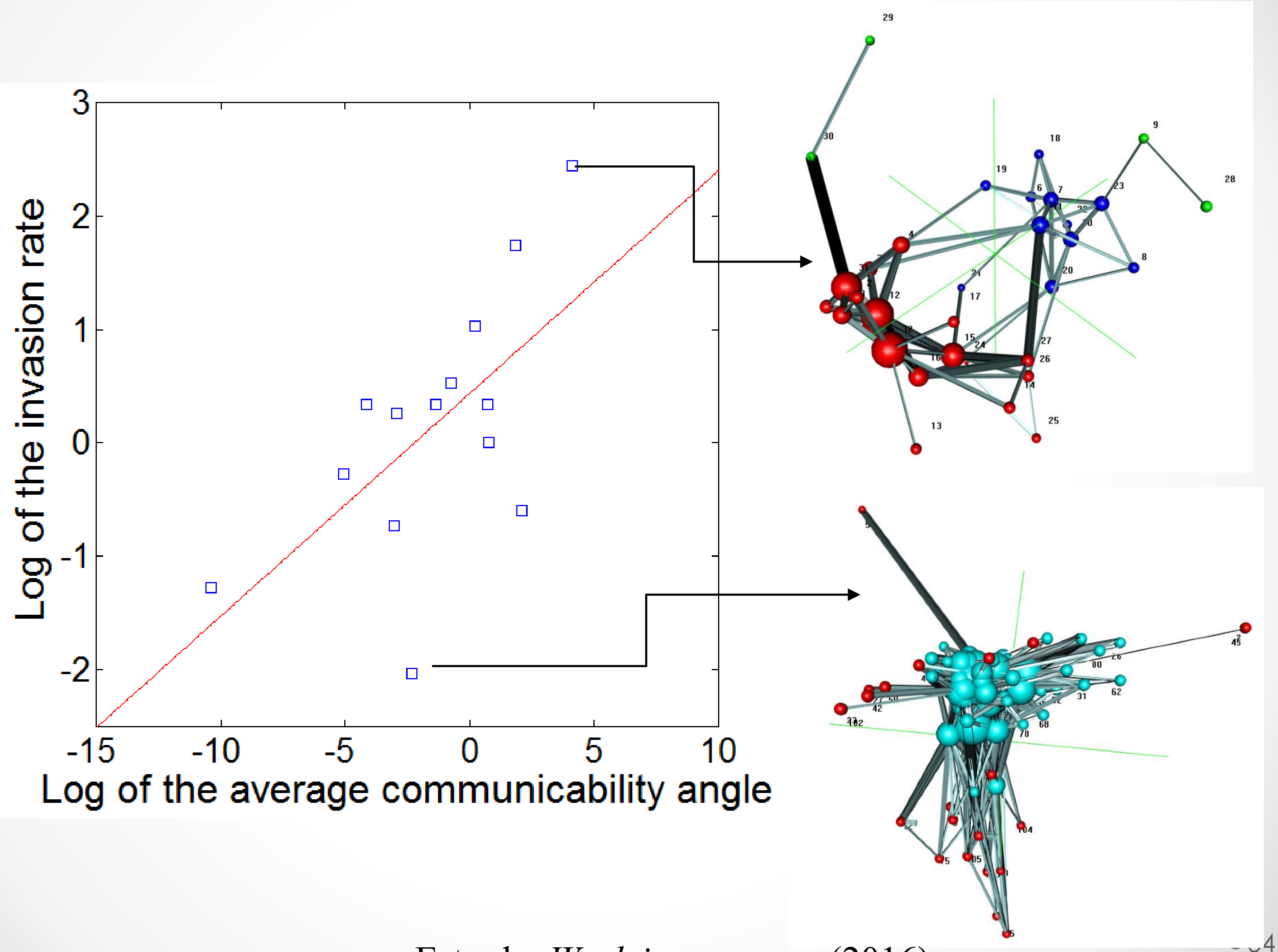
Multidimensional scaling





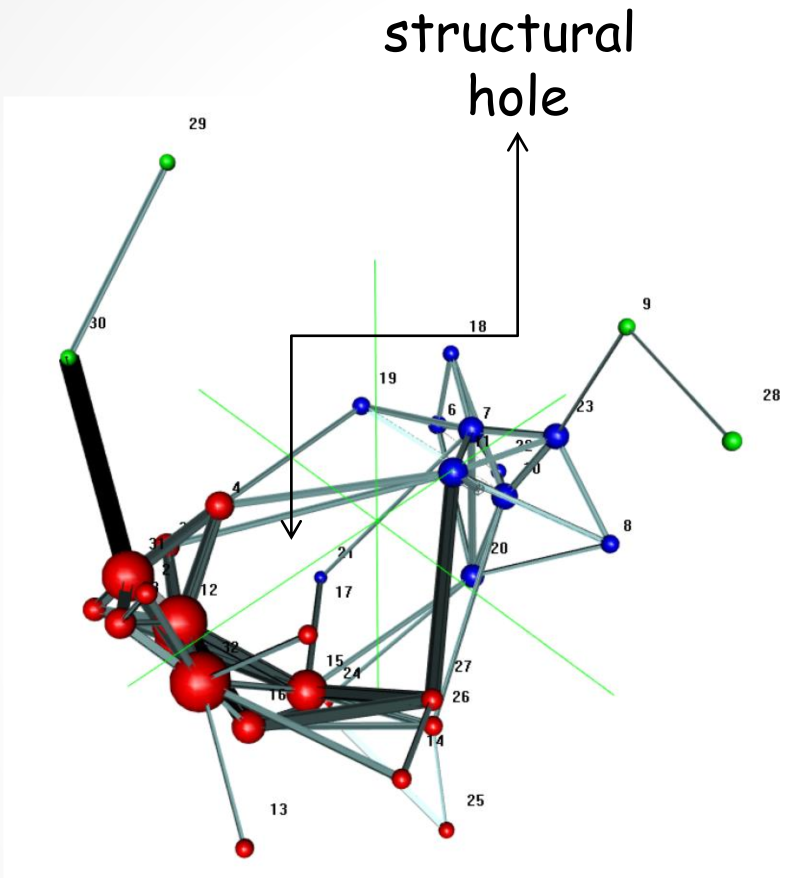
Estrada & Pereda: Work in progress (2016)

Network 'Spatial Efficiency'

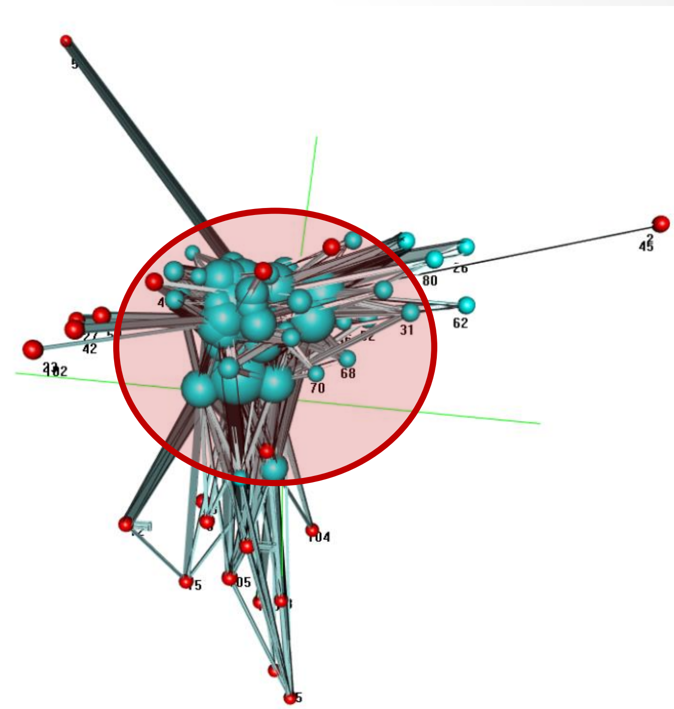


Estrada: *Work in progerss*, (2016).

Network 'Spatial Efficiency'



60.7°

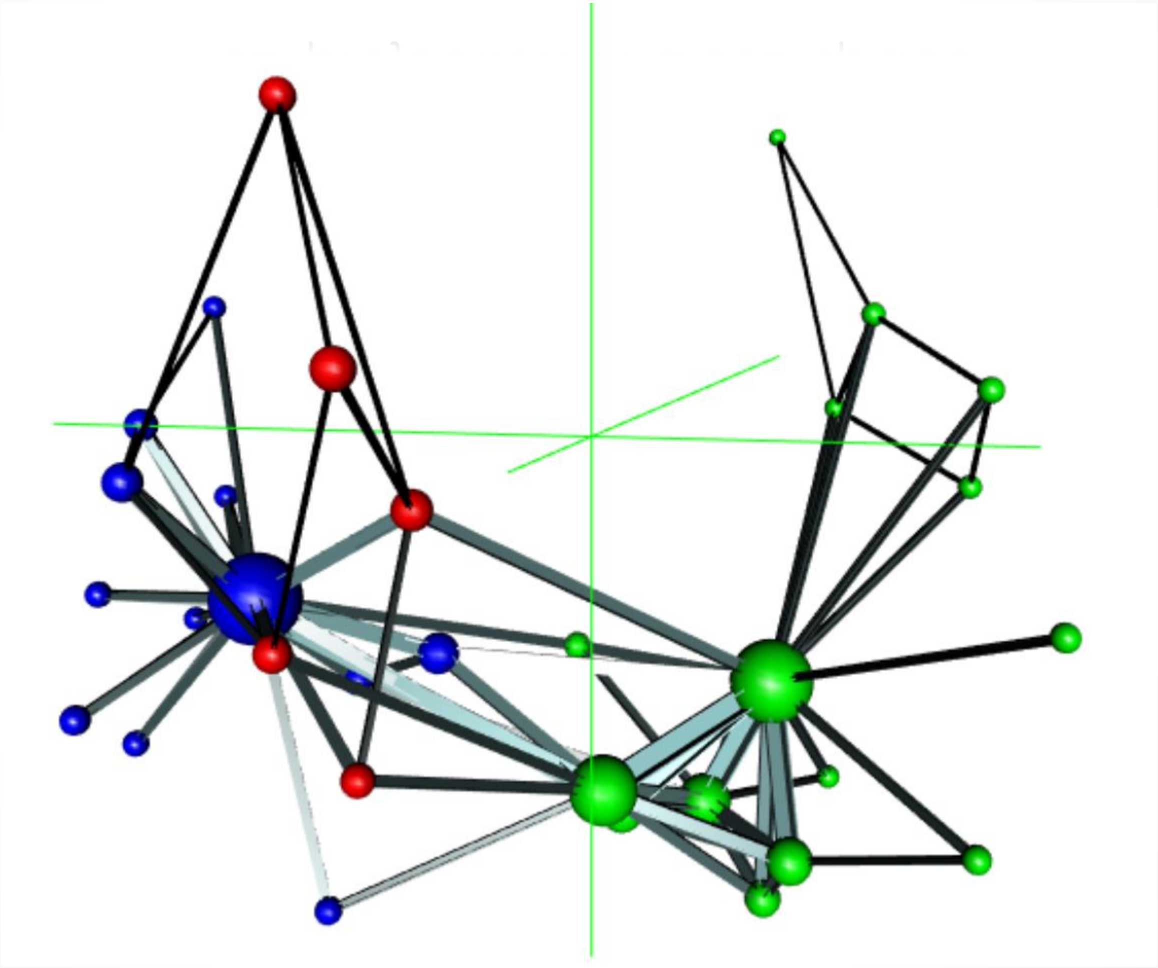


0.1°

Estrada: *Work in progerss*, (2016).

Reducing dimensionality

3D projection of a 33D space



Estrada & Pereda: Work in progress (2016)

θ_{pq} and Consensus Dynamics

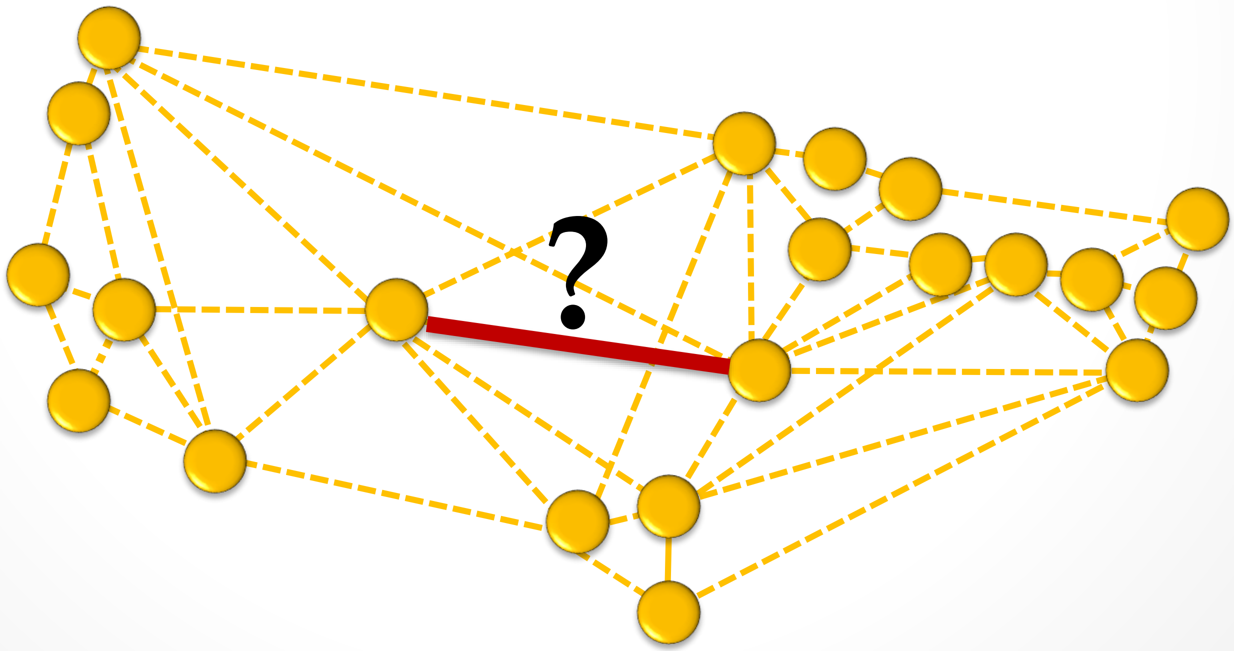
$$\dot{\vec{u}} = -L(G)\vec{u}(t),$$
$$\vec{u}(0) = \vec{u}_0$$

$$L(i,j) = \begin{cases} k_i & i = j \\ -1 & (i,j) \in E \\ 0 & (i,j) \notin E \end{cases}$$

Mesbahi, Egerstedt, *Graph Theoretic Methods in Multiagent Networks*, Pricenton Univ. Press, 2010, pp. 42-48.

θ_{pq} and Consensus Dynamics

Problem. *Given a network G, which are the most critical edges for a consensus (diffusion) dynamics. In other words, which are the edges whose removal do not disconnect the network but significantly increase the time for global consensus.*



Estrada, Vargas-Estrada & Ando: *Phys. Rev. E 92* **2015**, 052809.

θ_{pq} and Consensus Dynamics

Method:

Rank the edges according to a given centrality index:

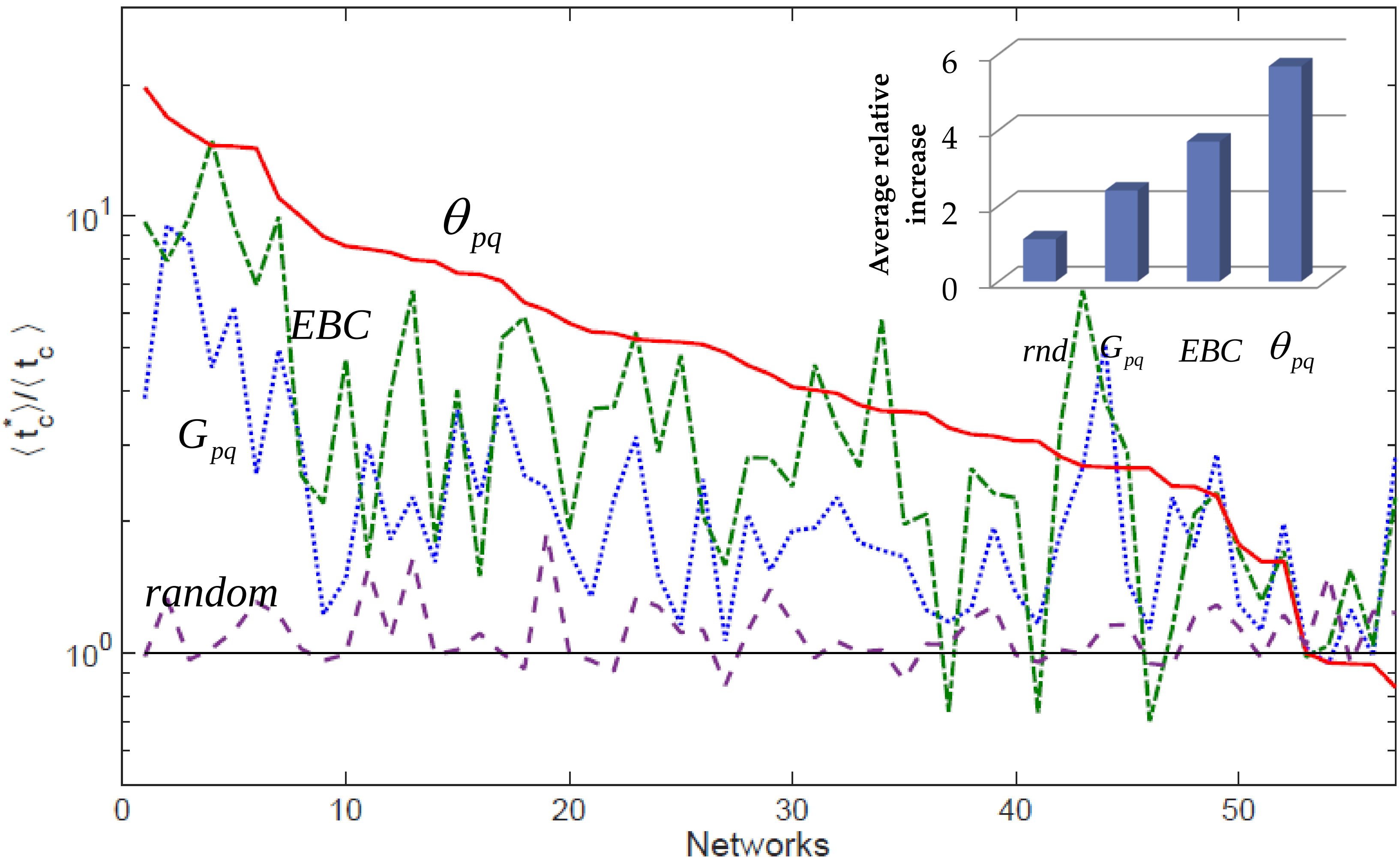
- ✓ *edge betweenness,*
- ✓ *average communicability of the vertices in the edge,*
- ✓ *communicability angle*

Remove 20% of the edges in the top of each ranking, guaranteeing the connectivity of the network.

Compare the consensus time with the original network and with the random removal of 20% of edges.

Estrada, Vargas-Estrada & Ando: *Phys. Rev. E 92* **2015**, 052809.

θ_{pq} and Consensus Dynamics



Estrada, Vargas-Estrada & Ando: *Phys. Rev. E 92* **2015**, 052809.

University of Strathclyde

Take home messages

- ✓ New mathematical and conceptual developments should be defined and studied for understanding the structural organisation of networks.
- ✓ The communicability distances and angles characterise the geometrical properties emerging from the communication among the nodes in a network.
- ✓ The geometric parameters of a network based on the communicability function capture important structural and dynamical characteristics of networks.

Thank you!