Short Course on Complex Networks and Point Processes with Applications

OXFORD, SEPTEMBER 11-12, 2017

Network Analytics. Traditional vs. Nontraditional Methods

Ernesto Estrada

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

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

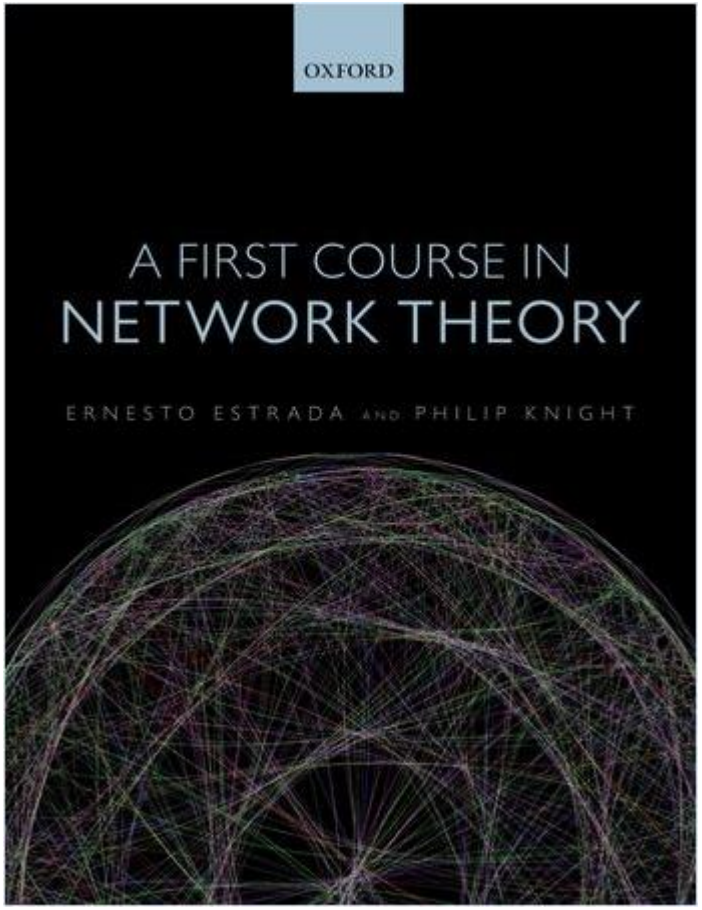
University of Strathclyde

Topics of the Lectures

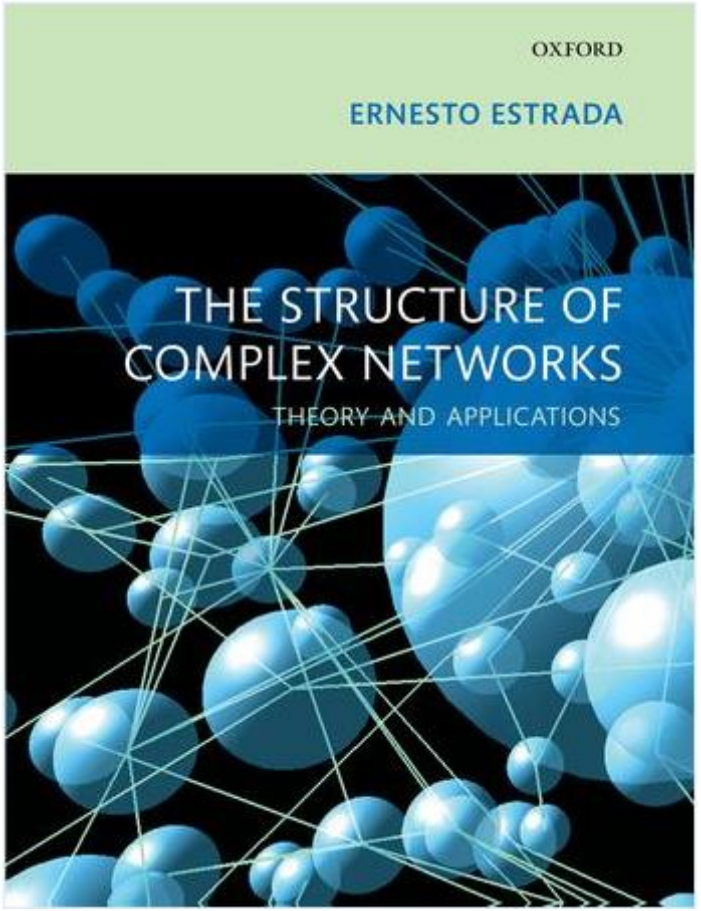
- Degree distributions
- Shortest paths vs. all-routes
- Clustering coefficients
- Geometry of networks
- Network bipartivity
- Degree-degree correlations

Topics of the Lectures

MAIN TEXT



AUXILIARY



University of Strathclyde

Motto of this course

Measure what is measurable, and make measurable what is not so.

Galileo Galilei

LECTURE 5: BASED ON CHAPTER 8 OF AUXILIARY BOOK

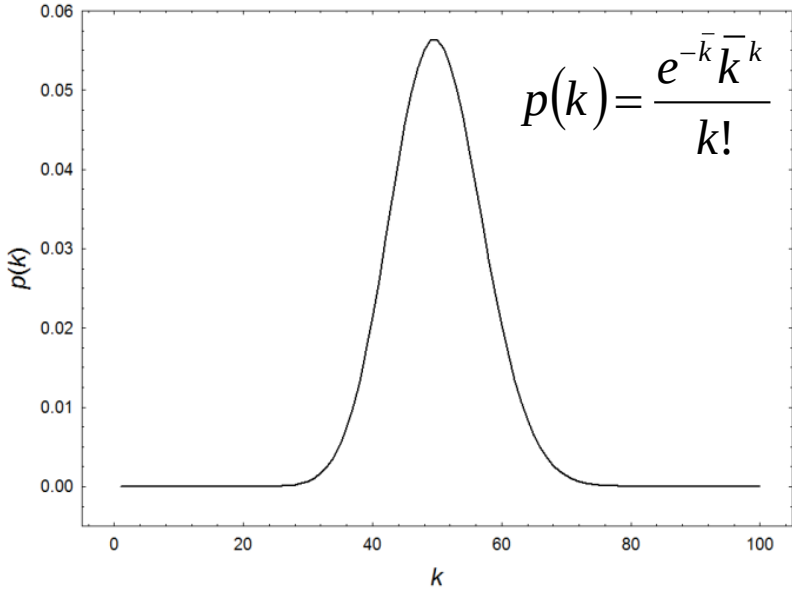
Degree Irregularity. Distributions vs. Algebraic Approaches

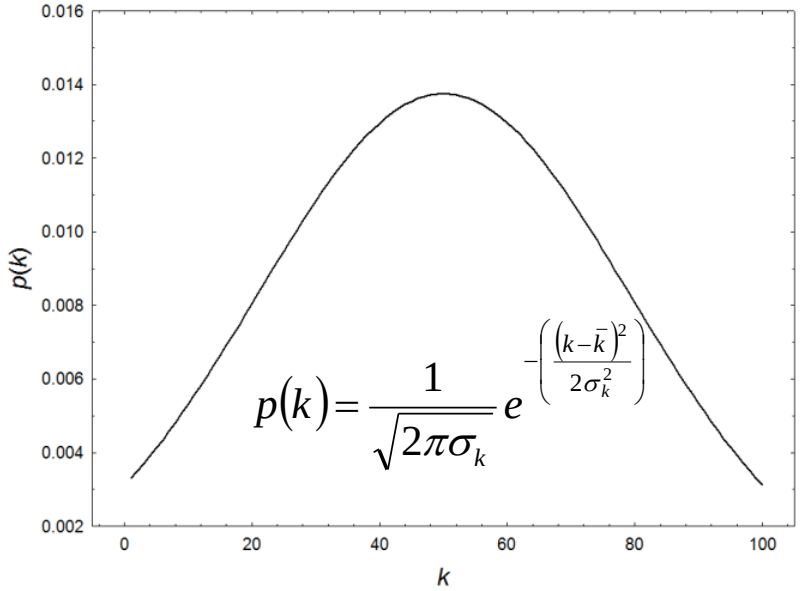
Ernesto Estrada

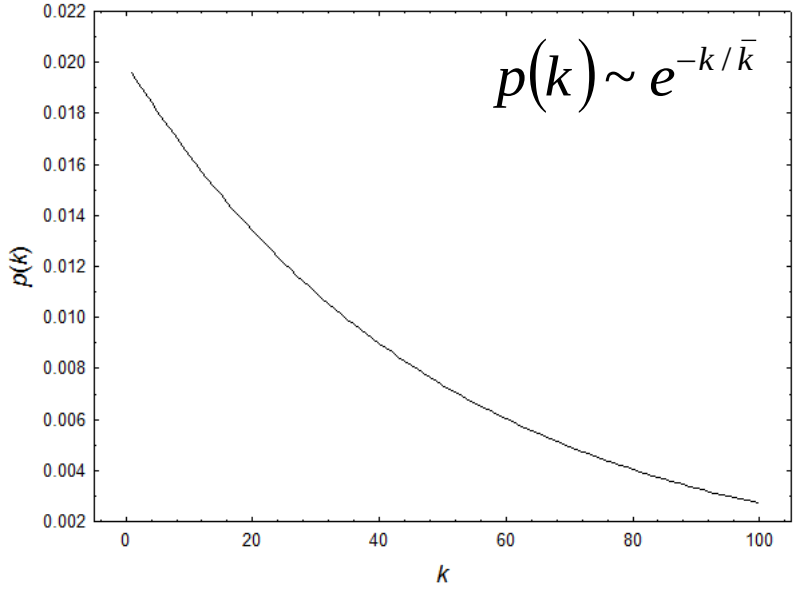
Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

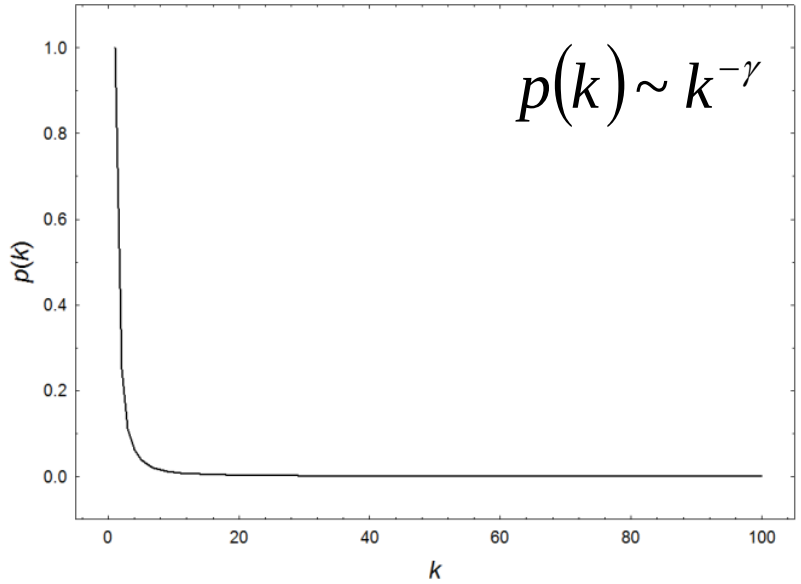
www.estradalab.org

Statistical Distributions



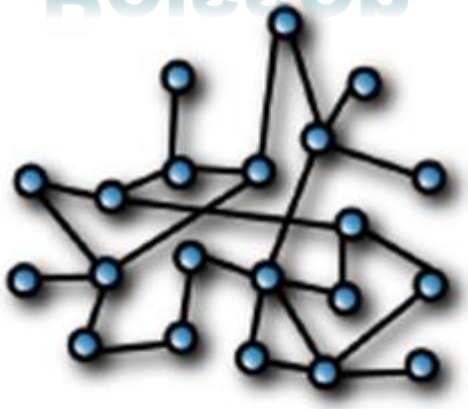


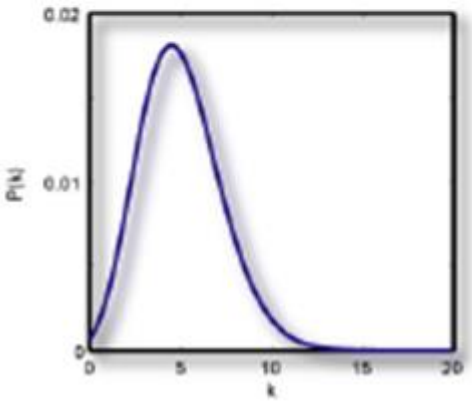




Degree Distributions

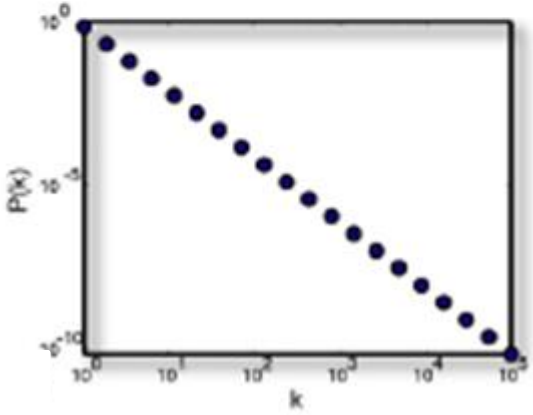
Poisson





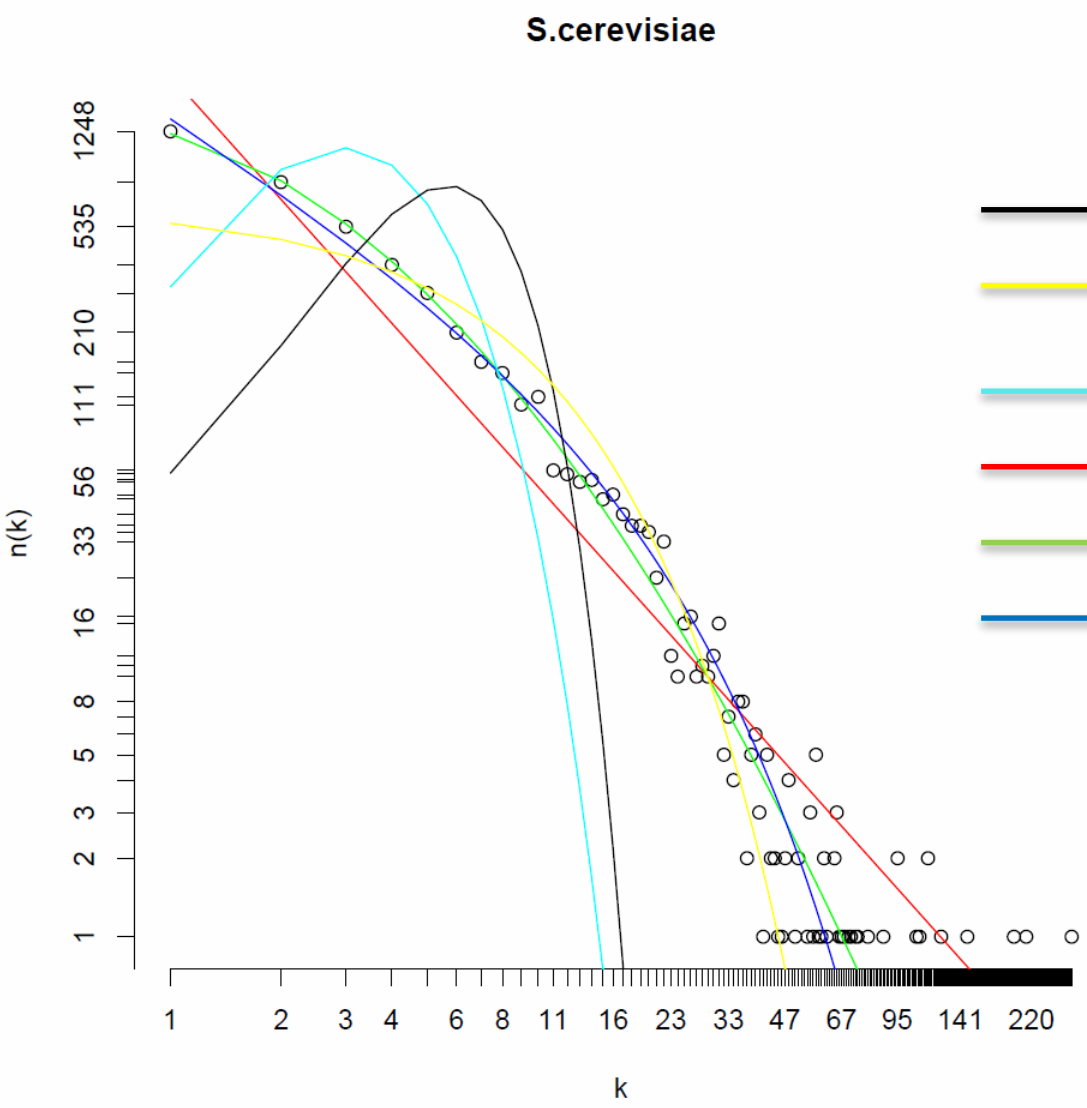
$$p(k) = \frac{e^{-\bar{k}}\bar{k}^{k}}{k!}$$

Power-law



$$p(k) \sim k^{-\gamma}$$

What is the best fit?



- Poisson
- Exponential
- Gamma
- Power-law
- Log-normal
- Stretched exponential

Stumpf & Ingram, *Europhys. Lett*. **71** (2005) 152.

The Zoo of Distributions

Bernoulli
Beta
Four-parameter Beta
Beta-Binomial
Beta-Geometric
Beta-Negative Bininomial
Binomial
Bradford
Burr
Cauchy
Chi
Chi-Square(d)
Cumulative Ascending
Cumulative Descending
Dagum
Delaporte
Discrete
Discrete Uniform
Error Function
m-Erlang

Generalized Error
Exponential
Extreme Value Max
Extreme Value Min
F
Fatigue Life
Gamma
Geometric
Generalized Logistic
Histogram
Hyperbolic Secant
Hypergeometric
Inverse Gaussian
Inverse Hypergeometric
Johnson Bounded
Johnson Unbounded

Kumaraswamy
Four-parameter Kumaraswamy
Laplace
Levy
Logarithmic (Series)
LogGamma
Logistic
LogLaplace
LogLogistic
Lognormal
LognormalB
LognormalE
Modified PERT
Negative Binomial
Normal
Ogive
Pareto

Pareto2
Pearson Type 5
Pearson Type 6
PERT
Poisson
Pólya
Rayleigh
Reciprocal
Relative
Slash
Split Triangle
Step Uniform
Student
Triangle
Uniform
Weibull

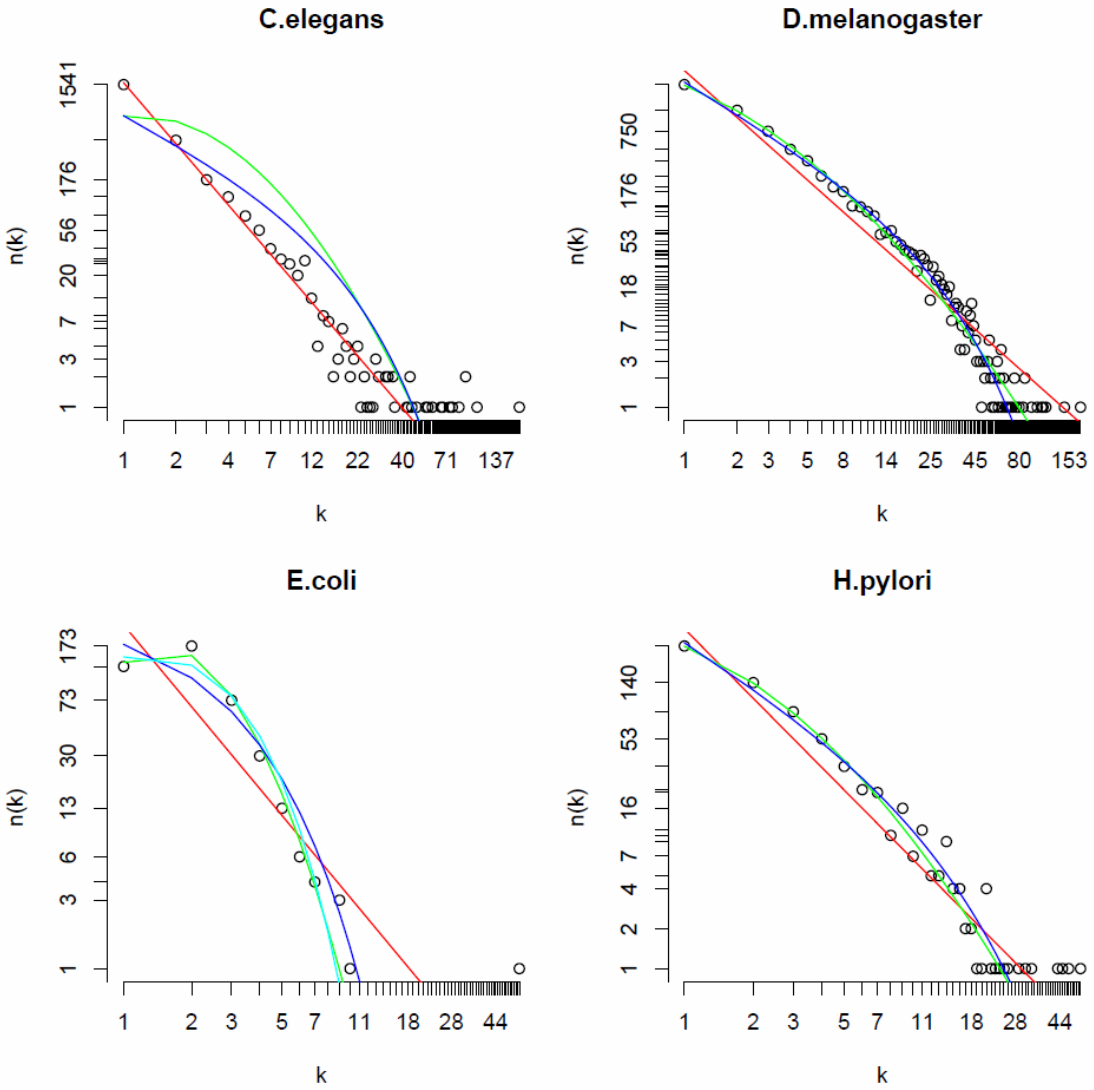
How to Compare?

Stretched exponential

$$p(k) \sim e^{-\alpha k / \bar{k}} k^{-\gamma}$$

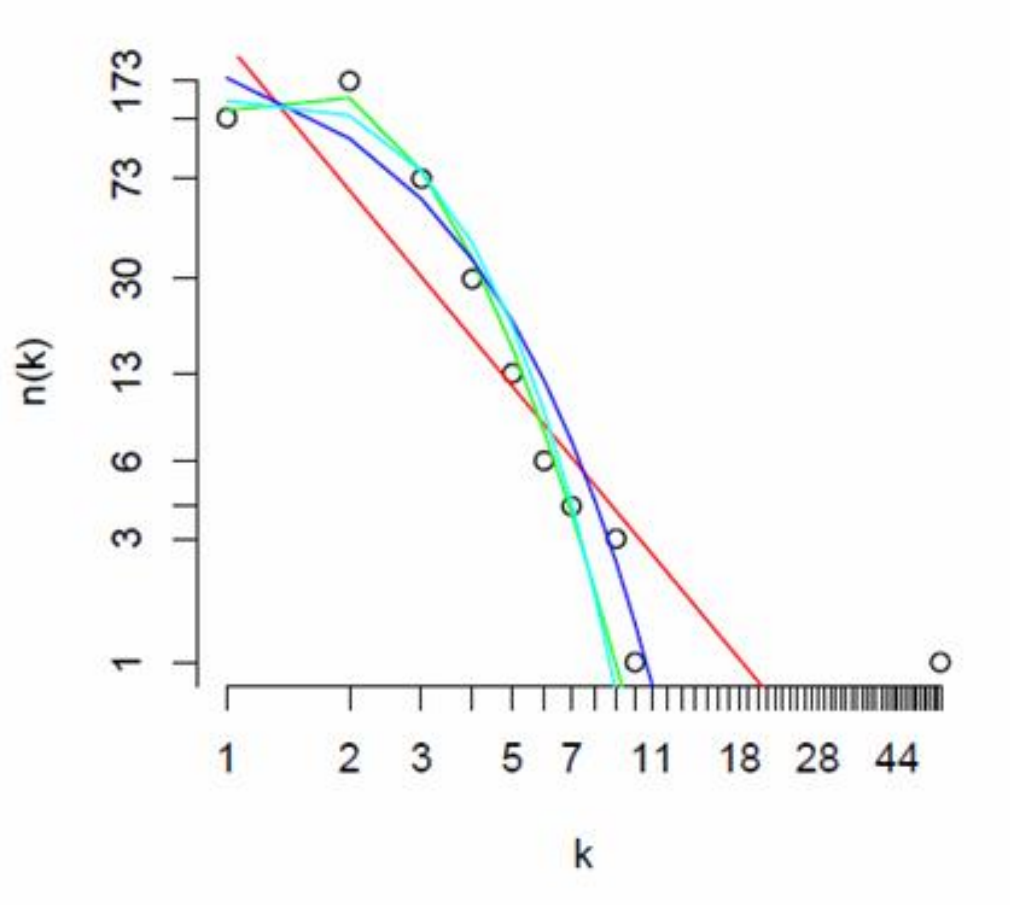
Log-normal

$$p(k) \sim \frac{e^{-\ln\left((k-\theta)/m\right)^2 / \left(2\sigma^2\right)}}{(k-\theta)\sigma\sqrt{2\pi}}$$



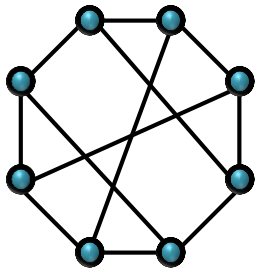
Stumpf & Ingram, *Europhys. Lett*. **71** (2005) 152.

How to Manage Scarce Data?

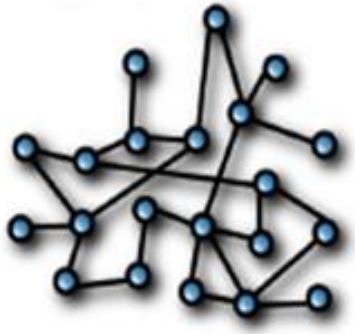


Is There an Order?

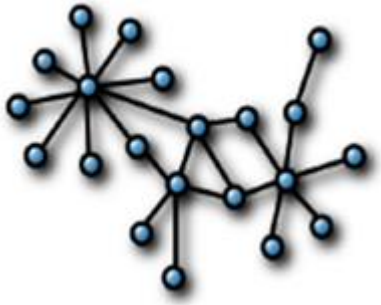
Regular



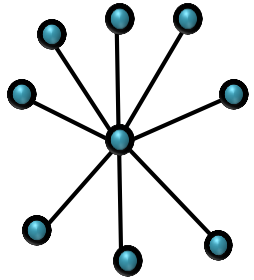
Poisson

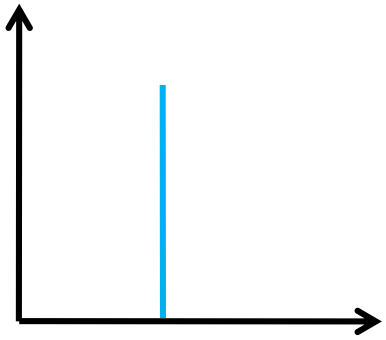


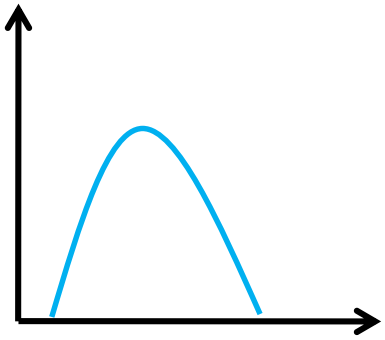
Fat-tail

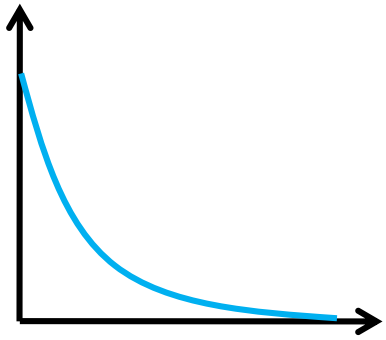


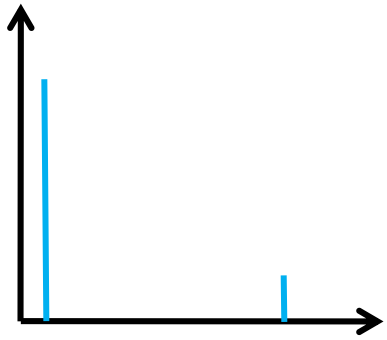
Star





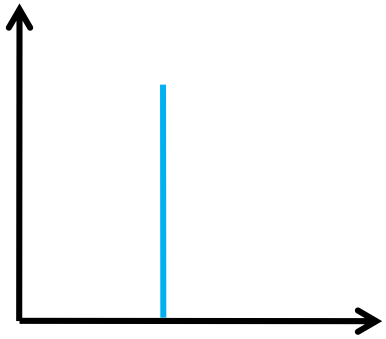


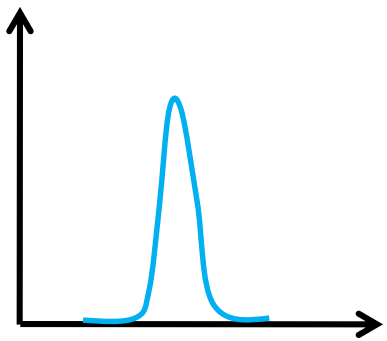


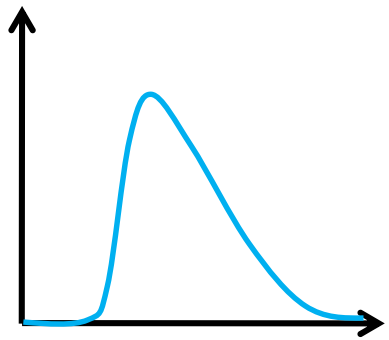


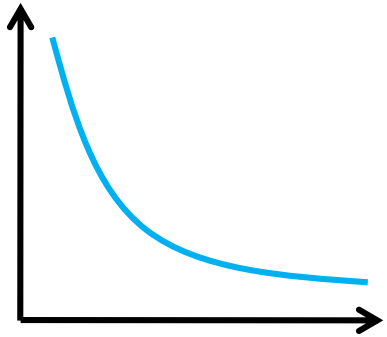
DEGREE HETEROGENEITY

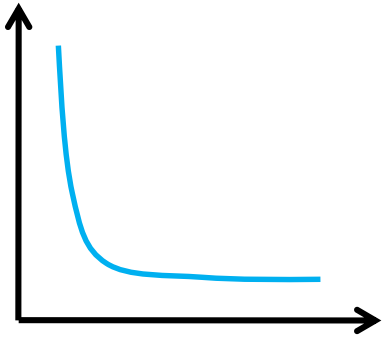
Is There an Order?

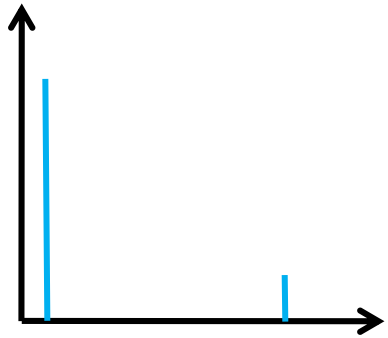












Distribution skewness?

Distribution skewness?

The qualitative interpretation of the skew is complicated. For example, a zero value indicates that the tails on both sides of the mean balance out, which is the case for a symmetric distribution, but is also true for an asymmetric distribution where the asymmetries even out, such as one tail being long but thin, and the other being short but fat. Further, in multimodal distributions and discrete distributions, skewness is also difficult to interpret. Importantly, the skewness does not determine the relationship of mean and median.

Wikipedia

Graph Theory Approach

Degree irregularity indices

- **Collatz-Sinogowitz Index (1957)**

$$CS(G) = \lambda_1 - \langle k \rangle$$

- **Bell Index (1992)**

$$Var(G) = \frac{1}{n}\sum_i k_i^2 - \left(\frac{1}{n}\sum_i k_i\right)^2$$

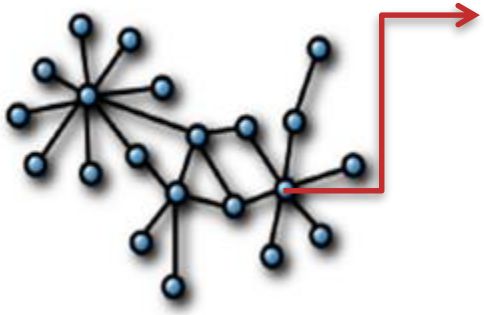
- **Albertson Index (1997)**

$$irr(G) = \sum_{(i,j)\in E} |k_i - k_j|$$

Unknown upper bounds

A New Index: Degree Skewness

probability of picking at random an incident edge to the node i



$$p_i = \frac{1}{k_i}$$

Non-standardised skewness: $\rho_{ij} = \langle p_{ij} \rangle - \hat{p}_{ij}$

Arithmetic mean

$$\langle p_{ij} \rangle = \frac{1}{2}\left(\frac{1}{k_i} + \frac{1}{k_j}\right)$$

geometric mean

$$\hat{p}_{ij} = \left(\frac{1}{\sqrt{k_i k_j}}\right)$$

Estrada: *Phys. Rev. E* **82**, 066102 (2010)

Degree Skewness

$$2\rho_{ij} = \left(\frac{1}{k_i} + \frac{1}{k_j}\right) - \frac{2}{\sqrt{k_i k_j}}$$

$$= \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}}\right)^2$$

Network degree skewness

$$\rho(G) = 2\sum_{(ij)\in E} \rho_{ij} = \sum_{(ij)\in E} \left(\frac{1}{\sqrt{k_i}} - \frac{1}{\sqrt{k_j}}\right)^2$$

The Laplacian Encounter

Graph Laplacian

$$L = K - A$$

$$L_{ij} = \begin{cases} k_i & \text{for } i = j, \\ -1 & \text{for } i \sim j, \\ 0 & \text{otherwise,} \end{cases}$$

Normalised Graph Laplacian

$$\mathcal{L} = K^{-1/2} \cdot L \cdot K^{-1/2}$$
$$= I - K^{-1/2} \cdot A \cdot K^{-1/2}$$

$$\mathcal{L}_{ij} = \begin{cases} 1 & \text{for } i = j, \\ -\dfrac{1}{\sqrt{k_i k_j}} & \text{for } i \sim j, \\ 0 & \text{otherwise,} \end{cases}$$

The Laplacian Encounter

Network degree skewness

$$\rho(G) = \vec{1} \cdot \mathcal{L} \cdot \vec{1}^{T}$$

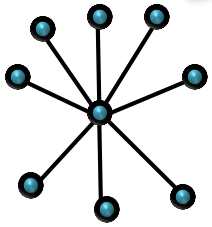
$$= n - 2 \sum_{(i,j) \in E} \left(k_i k_j \right)^{-1/2}$$

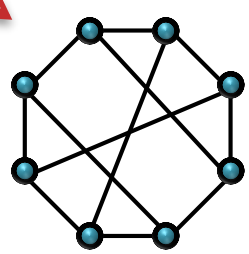
$$= n - 2\chi(G).$$

Randić index

A Nice Normalisation

$$\sqrt{n-1} \le \chi(G) \le \frac{n}{2}$$





Normalised network degree skewness

$$\tilde{\rho}(G) = \frac{n - 2\chi(G)}{n - 2\sqrt{n-1}}$$

$$1 \ge \tilde{\rho}(G) \ge 0$$

Back to the Normalised Laplacian

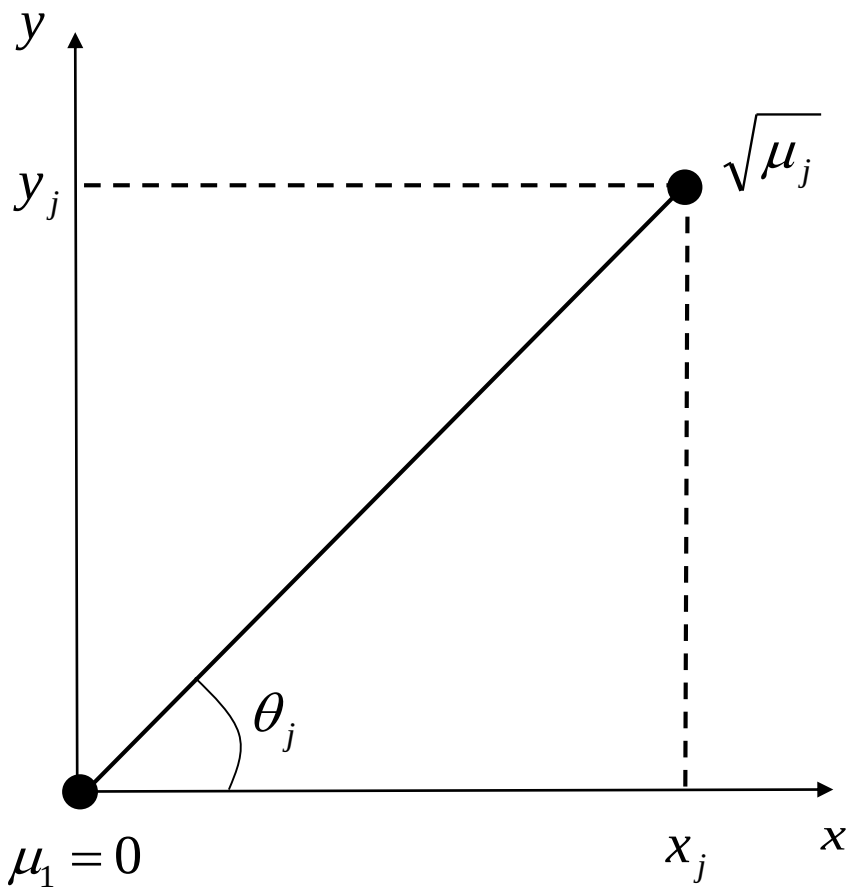
$$\rho(G) = \vec{1}^T \cdot \mathfrak{L} \cdot \vec{1} = n \sum_j \mu_j \cos^2 \theta_j$$

$$0 = \mu_1 < \mu_2 \leq \cdots \leq \mu_n$$

$$\cos \theta_j = \frac{1^T \cdot \vec{\varphi}_j}{\|1\| \|\vec{\varphi}_j\|} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_j(i)$$

$$\rho(G) = \vec{1}^T \cdot \mathfrak{L} \cdot \vec{1} = \sum_j \mu_j \left(\sum_i \varphi_j(i) \right)^2$$

Spectral Representation



$$x_j = \sqrt{\mu_{j>1}} \cos\theta_j$$

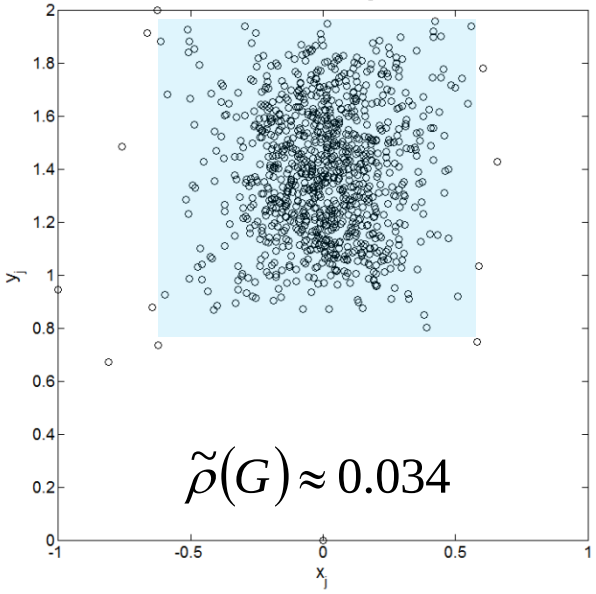
$$y_j = \sqrt{\mu_{j>1}} \sin\theta_j$$

$$\tilde{\rho}(G) = \frac{n}{n - 2\sqrt{n-1}} \sum_{j=1}^{n} x_j^2$$

Results

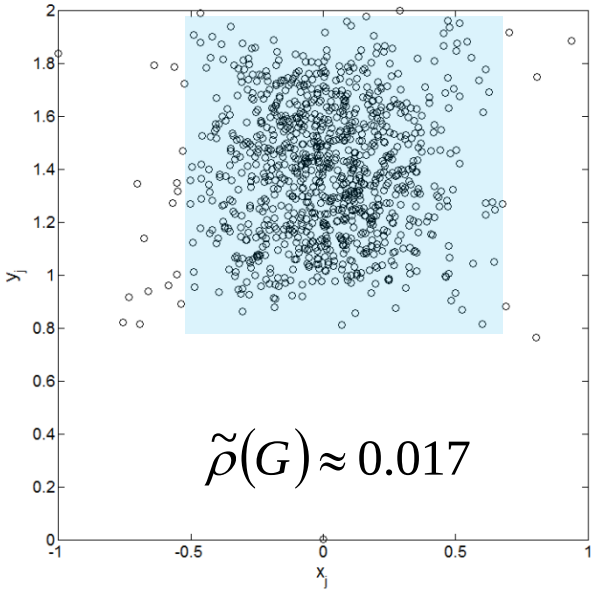
Erdös-Rényi Random Graphs $G_{n,p}$

$\bar{k} = 8$



$\tilde{\rho}(G) \approx 0.034$

$\bar{k} = 16$

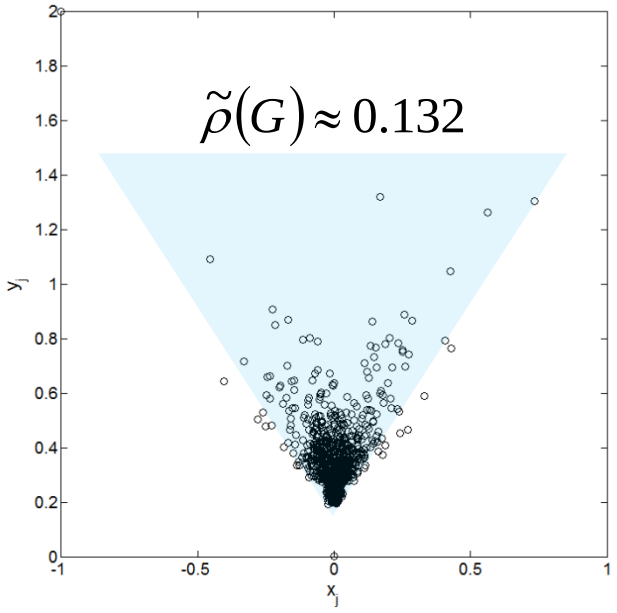


$\tilde{\rho}(G) \approx 0.017$

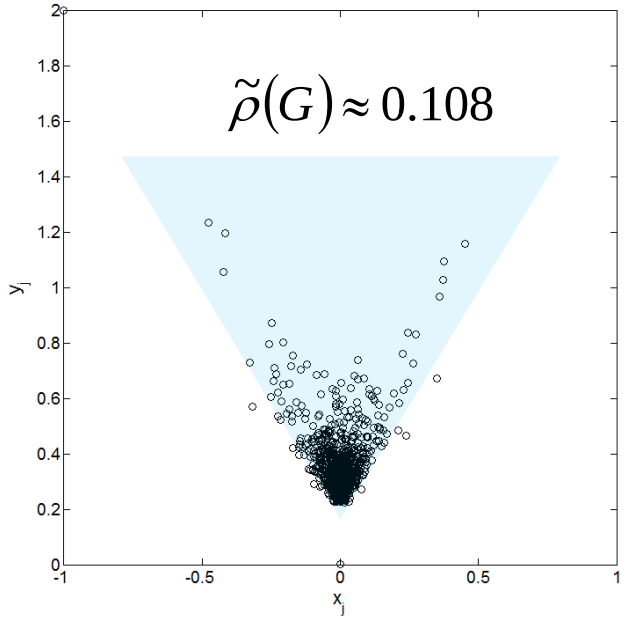
Barabási-Albert Random Graphs

$$\rho(BA) \approx \left(\frac{0.27 + \langle k \rangle^{-1.18}}{2.27 + \langle k \rangle^{-1.18}} \right) \left(\frac{n}{n - 2\sqrt{n-1}} \right)$$

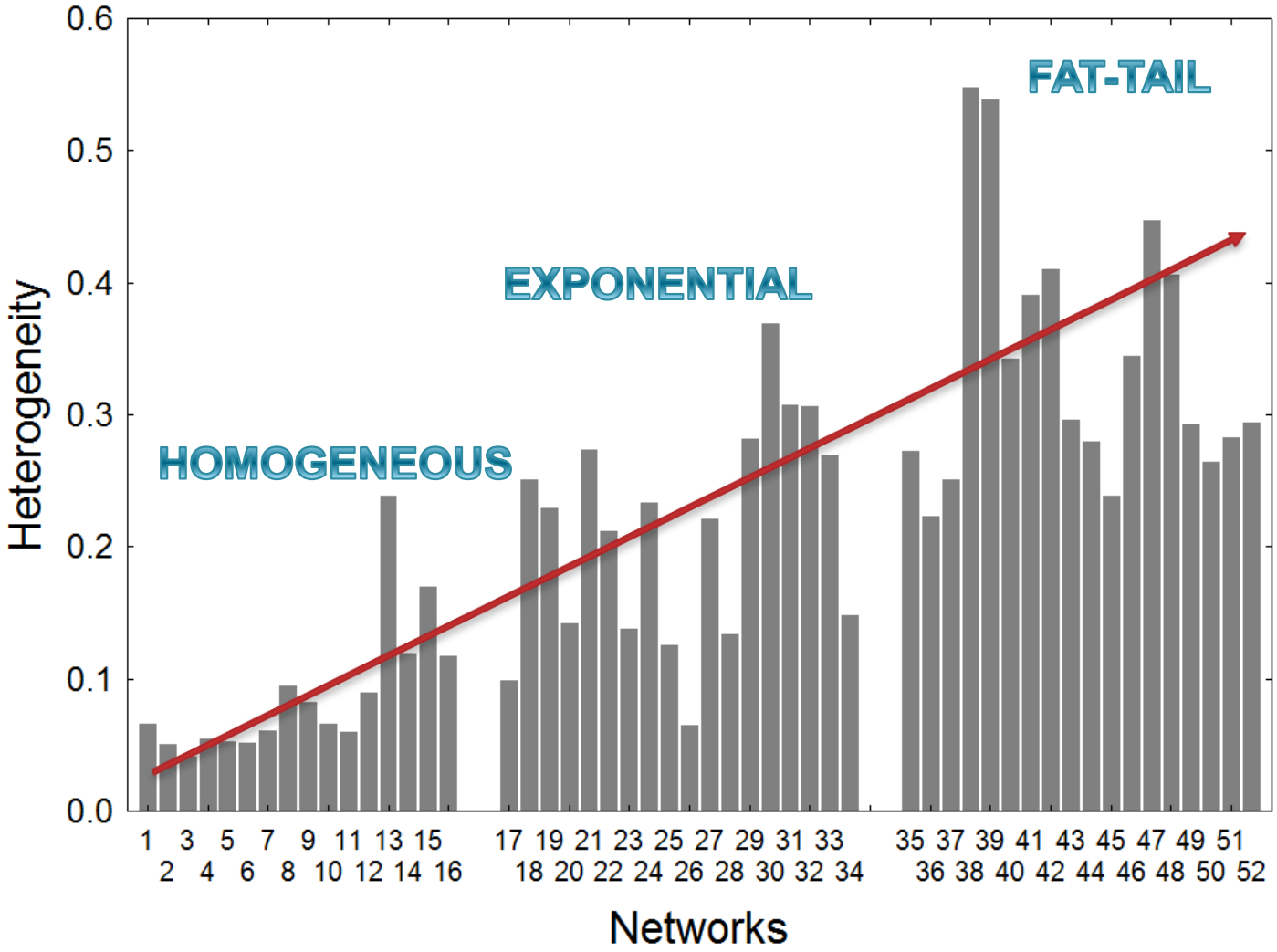
$\bar{k} = 8$



$\bar{k} = 16$



Real-World Networks

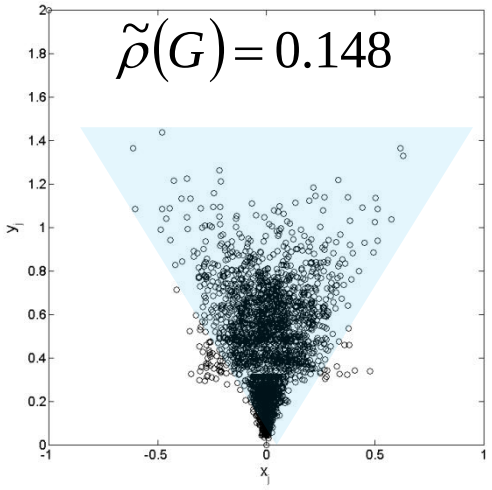


But the devil is in the details

Real-World Networks

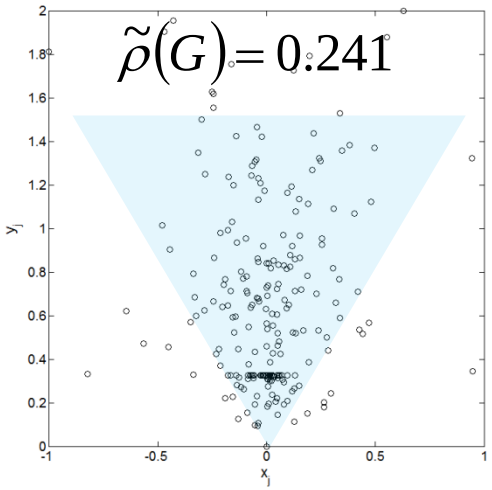
D. melanogaster

$\tilde{\rho}(G) = 0.148$



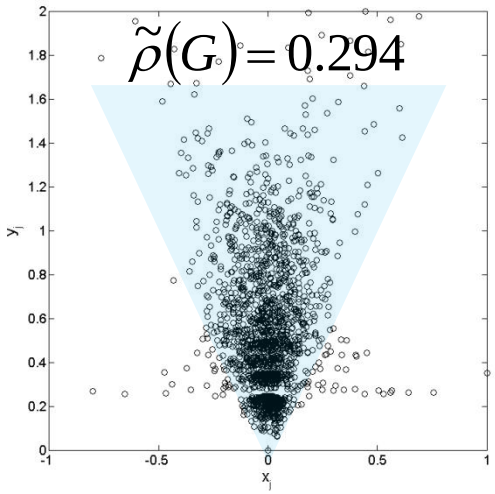
E. coli

$\tilde{\rho}(G) = 0.241$



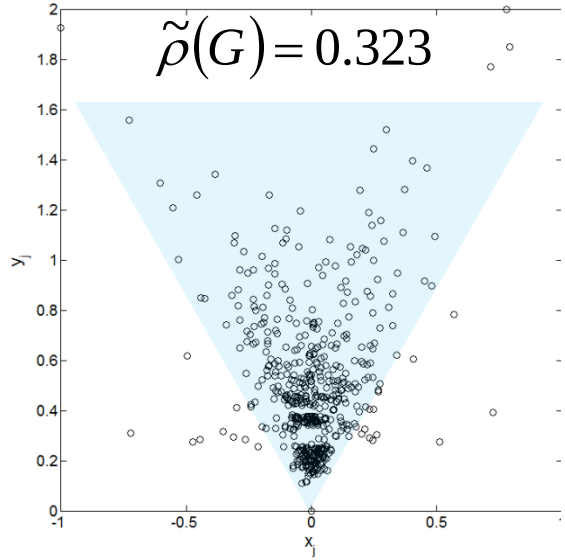
S. cereviciae

$\tilde{\rho}(G) = 0.294$



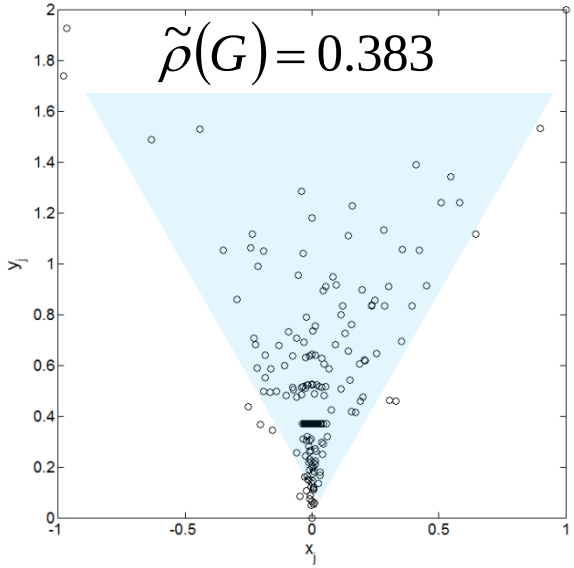
H. pylori

$\tilde{\rho}(G) = 0.323$



C. elegans

$\tilde{\rho}(G) = 0.383$



Stretched exponential

$$p(k) \sim e^{-\alpha k / \bar{k}} k^{-\gamma}$$

Log-normal

$$p(k) \sim \frac{e^{-\ln\left((k-\theta)/m\right)^2 / \left(2\sigma^2\right)}}{(k-\theta)\sigma\sqrt{2\pi}}$$

LECTURE 8: BASED ON CHAPTERS 11 & 19

Shortest Path vs. All-Routes Communication in Networks

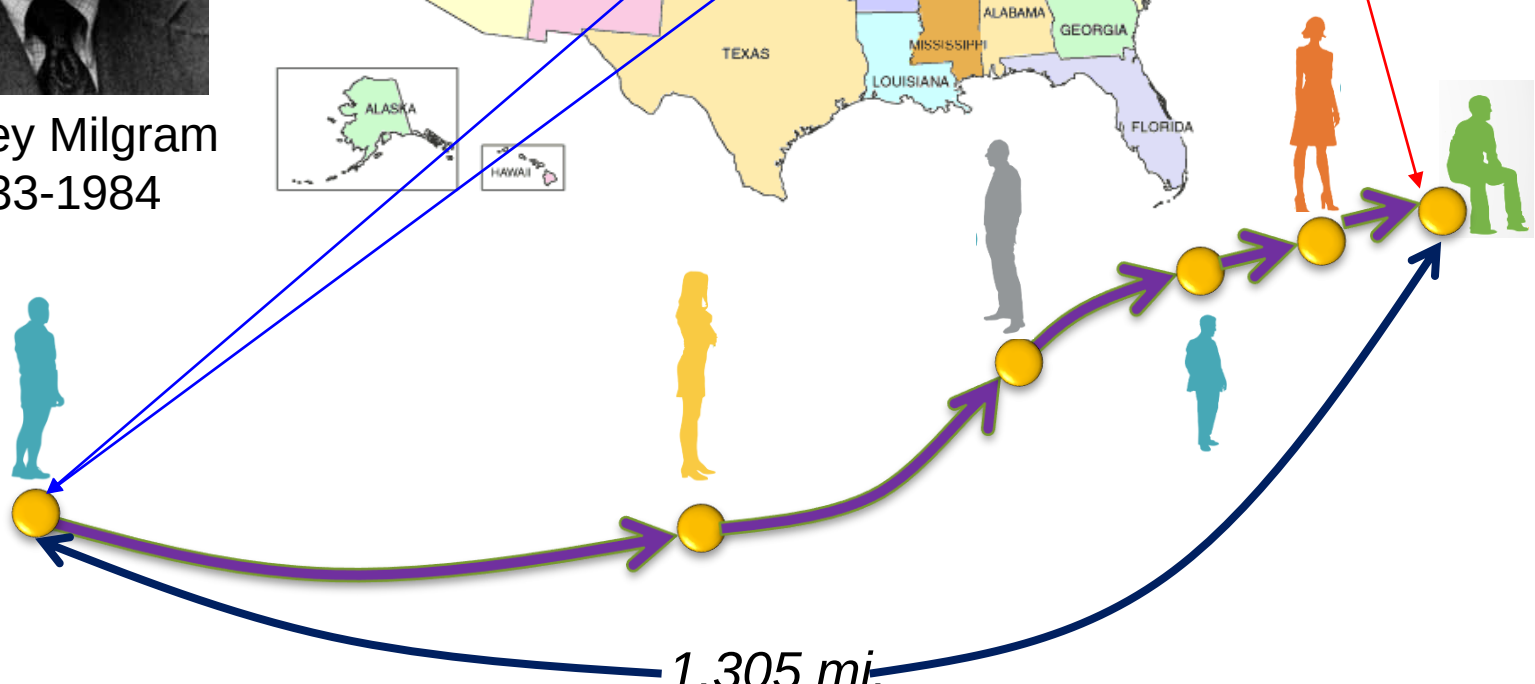
Ernesto Estrada

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

The 'Small-World' Phenomenon

Stanley Milgram
1933-1984

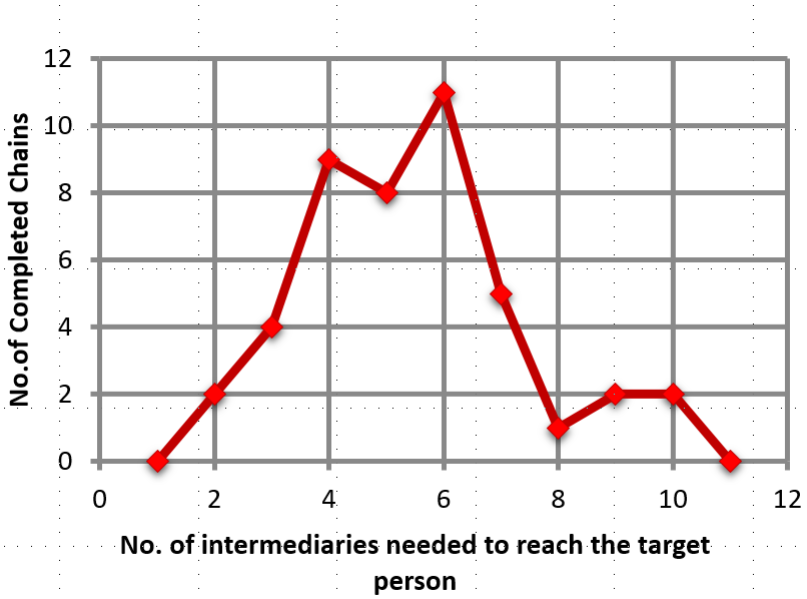


1,305 mi.

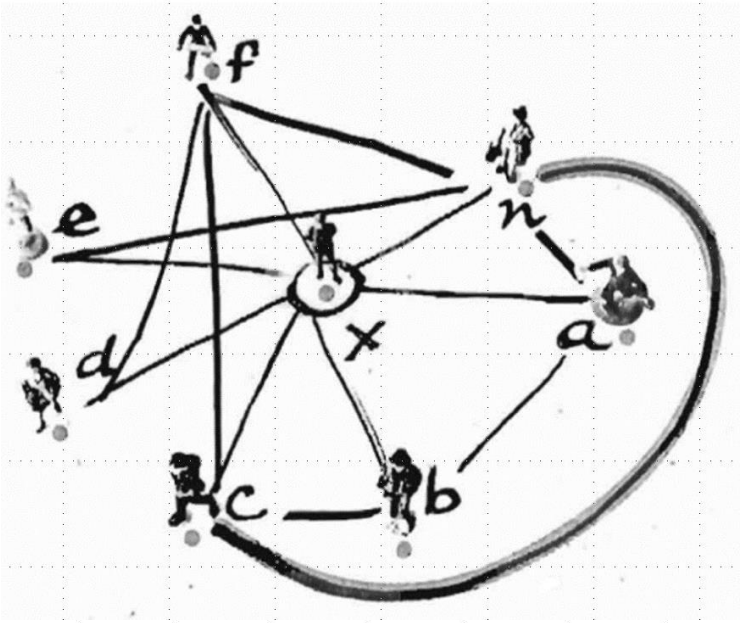
(see Chapter 11)

'Small-World' Phenomenon

Six Degrees of Separation



Imbreeding

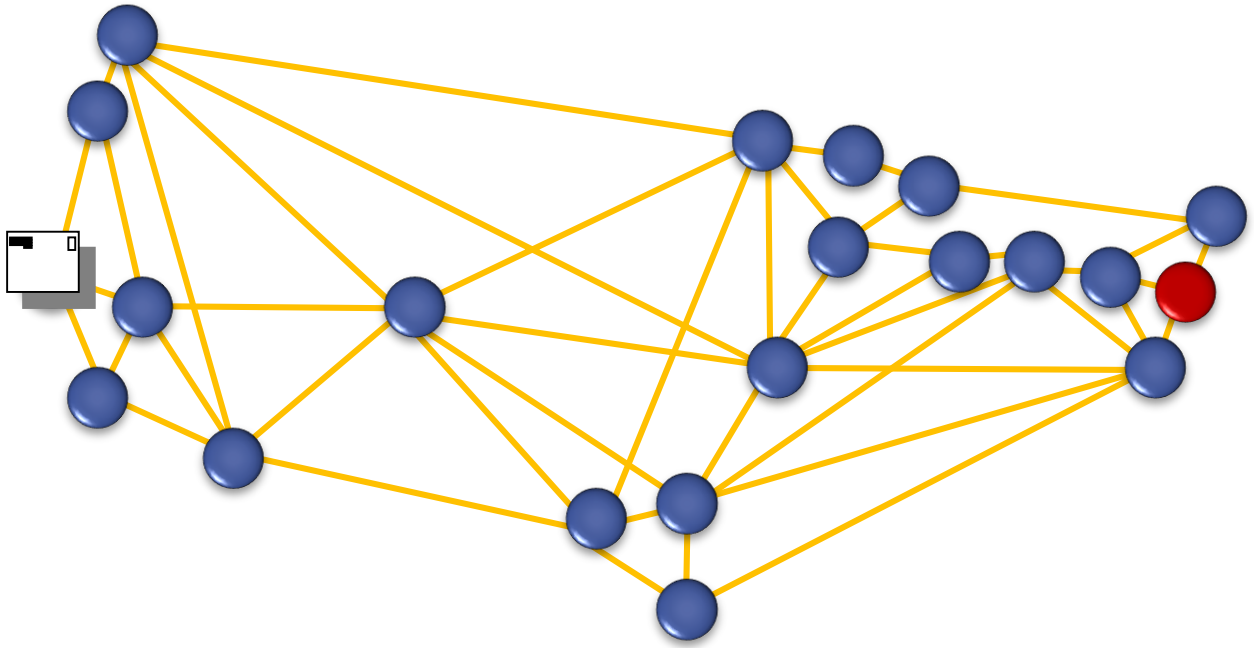


With group inbreeding, X's acquaintances feed back into his own circle, normally eliminating new contacts.

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

Shortest Path Distance

A path of length l is any sequence of distinct nodes $v_1, v_2, \ldots, v_l$ such as for each $i = 1, 2, \ldots, l$ there is a link from v_i to v_{i+1}.



$$d_{ij} = 5$$

Shortest Path Distance

(see Chapter 11)

How to quantify 'imbreeding'?

Steven Strogatz

Definition 1. *Let* $G=(V,E)$ *be a connected, simple, undirected network, and let* $i \in V$. *Then, the clustering coefficient of the node* i *is given by*

$$C_i = \frac{\text{number of transitive relations of node } i}{\text{total number of possible transitive relations of node } i}$$

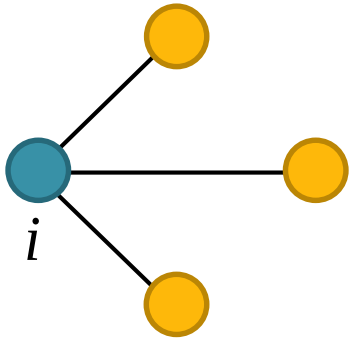
Duncan Watts

Mathematically,

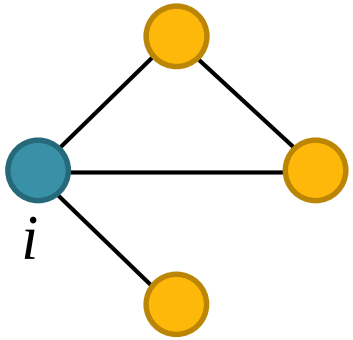
$$C_i = \frac{t_i}{k_i(k_i-1)/2} = \frac{2t_i}{k_i(k_i-1)}.$$

How to quantify 'imbreeding'?

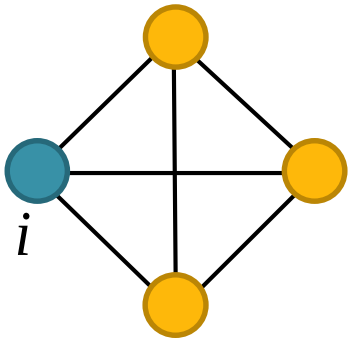
Example 1



$C_i = 0$



$C_i = 1/3$



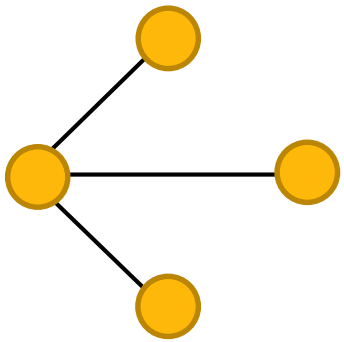
$C_i = 1$

Watts-Strogatz (WS) Clustering

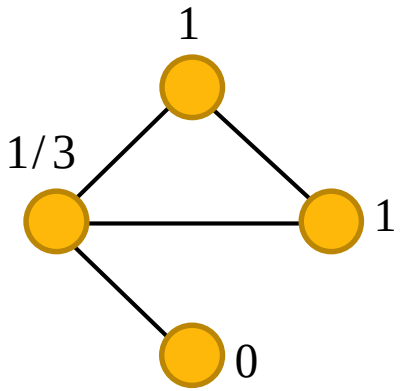
Average WS clustering coefficient of a network

$$\overline{C} = \frac{1}{n}\sum_i C_i$$

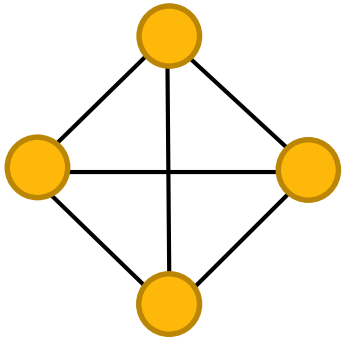
Example 2



$\overline{C} = 0$

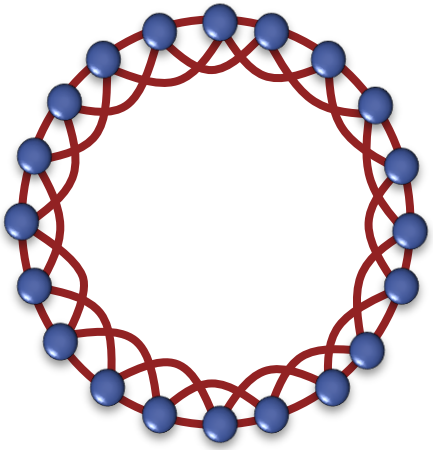
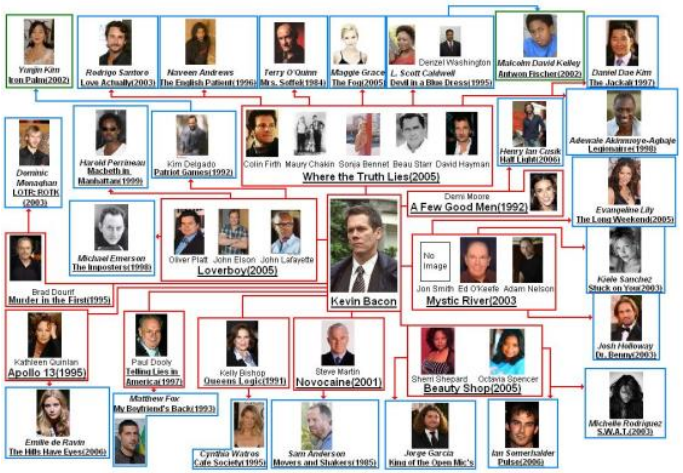
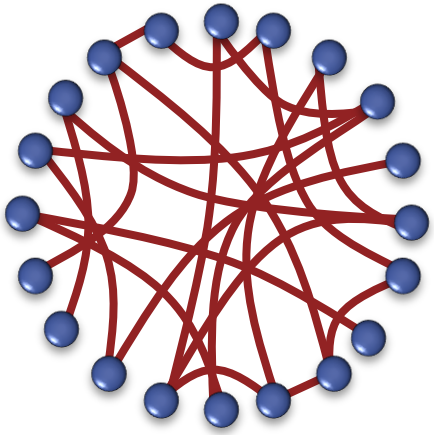


$\overline{C} = 7/12$



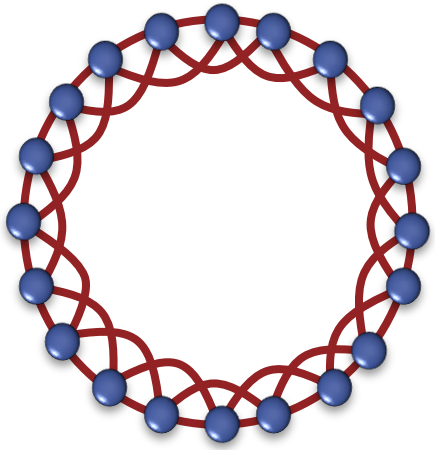
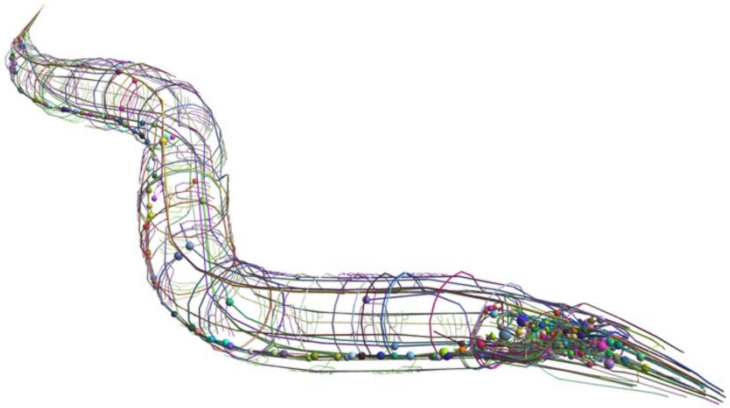
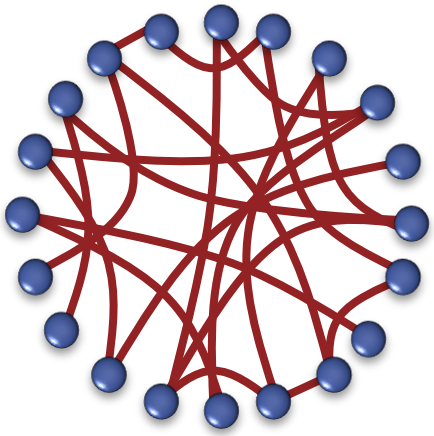
$\overline{C} = 1$

Shortest Path Approaches

	Regular	Hollywood Actors	Random
			
$\bar{d}$	1847	3.65	2.99
$\bar{C}$	0.74	0.79	0.00027

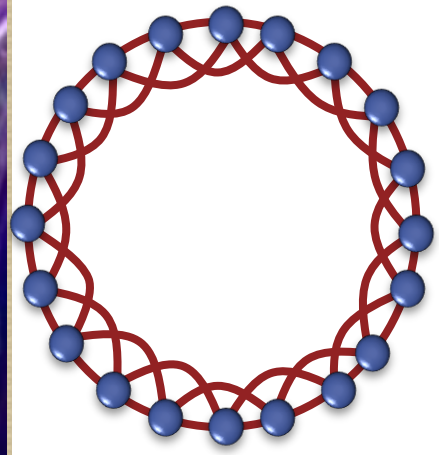
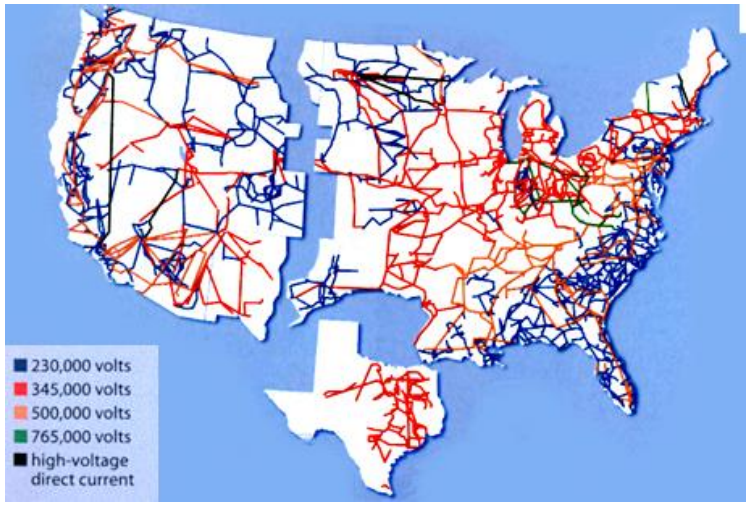
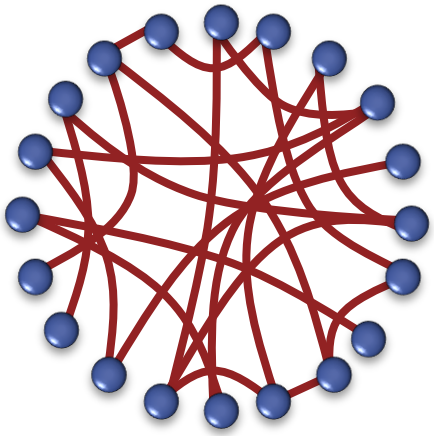
(see Chapter 11)

Shortest Path Approaches

Regular	Neural network of *C. elegans*	Random
		
$\bar{d}$ 10.5	2.65	2.25
$\bar{C}$ 0.69	0.28	0.05

(see Chapter 11)

Shortest Path Approaches

Regular	Power Grid of Western US	Random
		

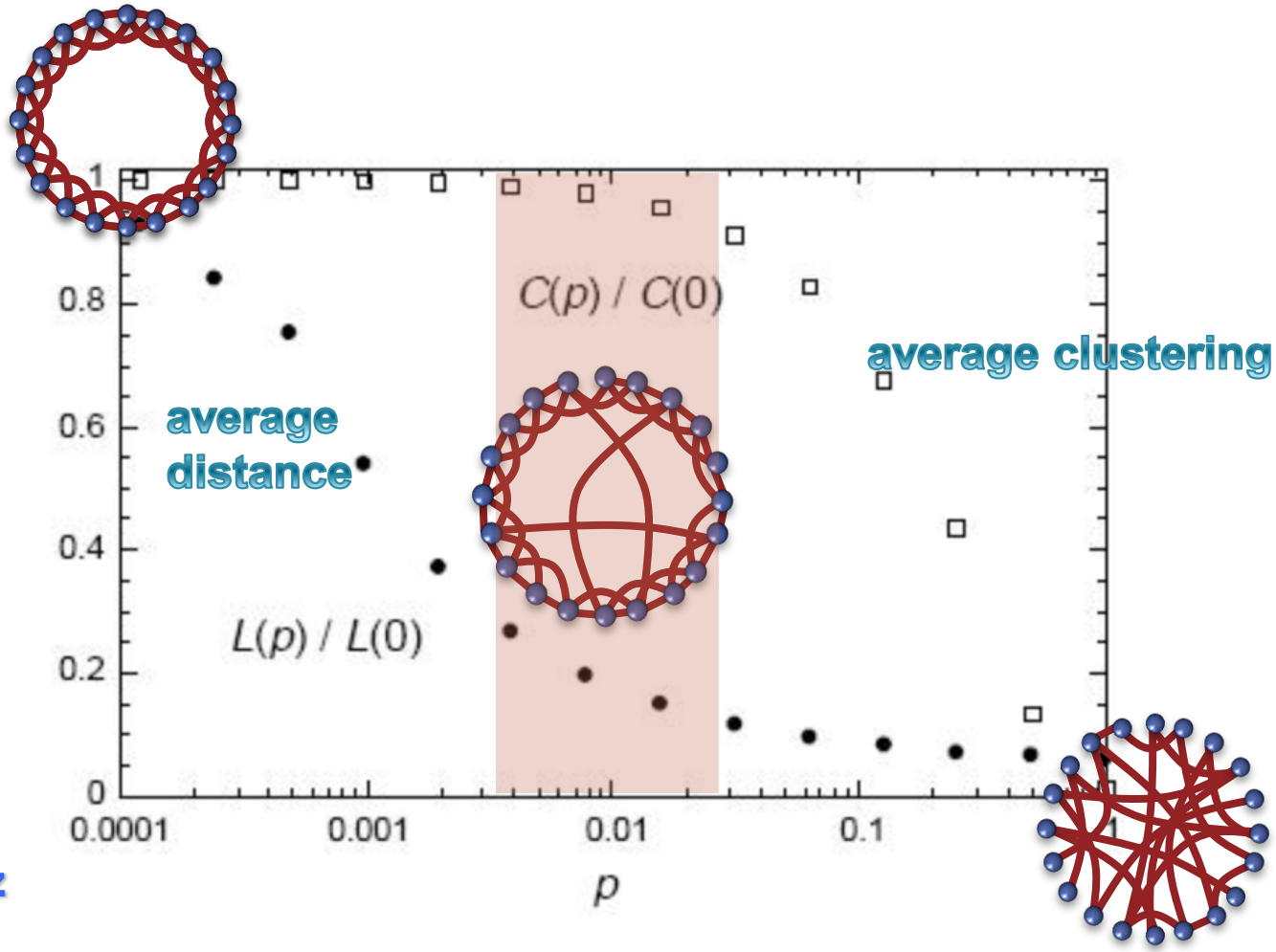
	Regular	Power Grid of Western US	Random
$\bar{d}$	926	18.7	12.4
$\bar{C}$	0.3	0.08	0.0005

(see Chapter 11)

Shortest Path Approaches

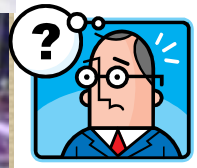
D. J. Watts

S. H. Strogatz

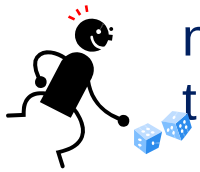


(see Chapter 11)

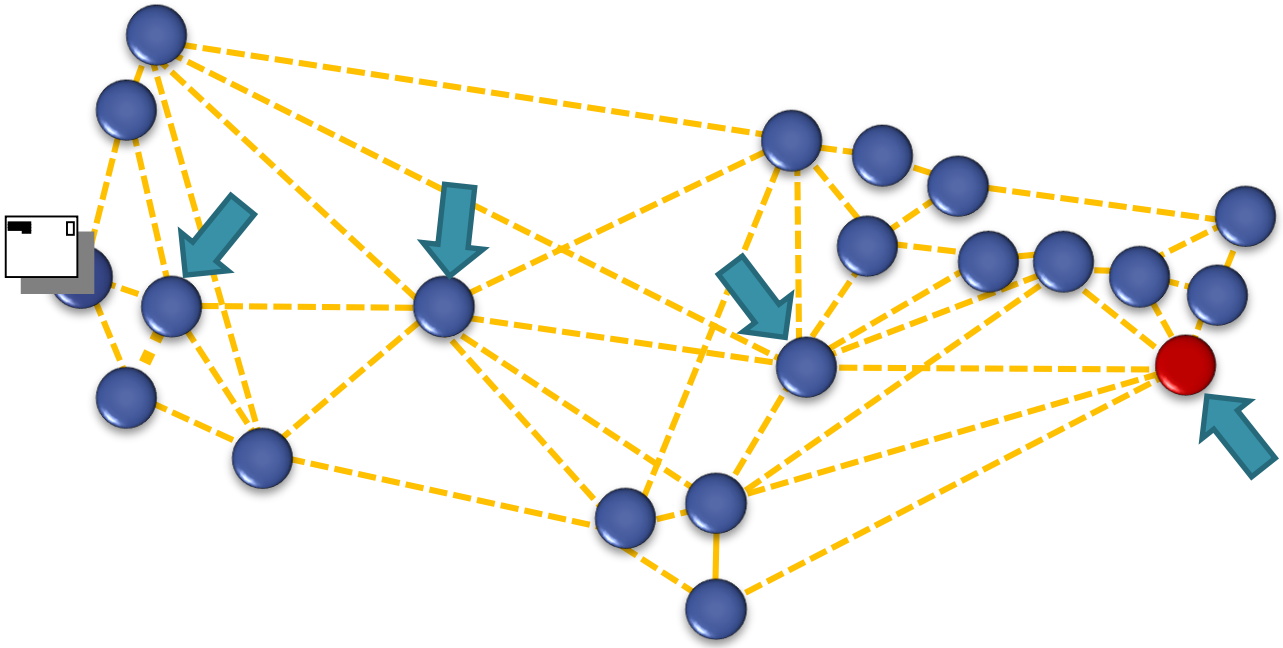
Shortest Path Approaches



The sender does not know the global structure of the network!



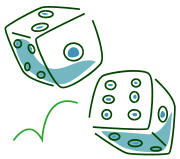
There is a high probability that the shortest path connecting nodes *p* and *q* goes through the most connected nodes of the network.



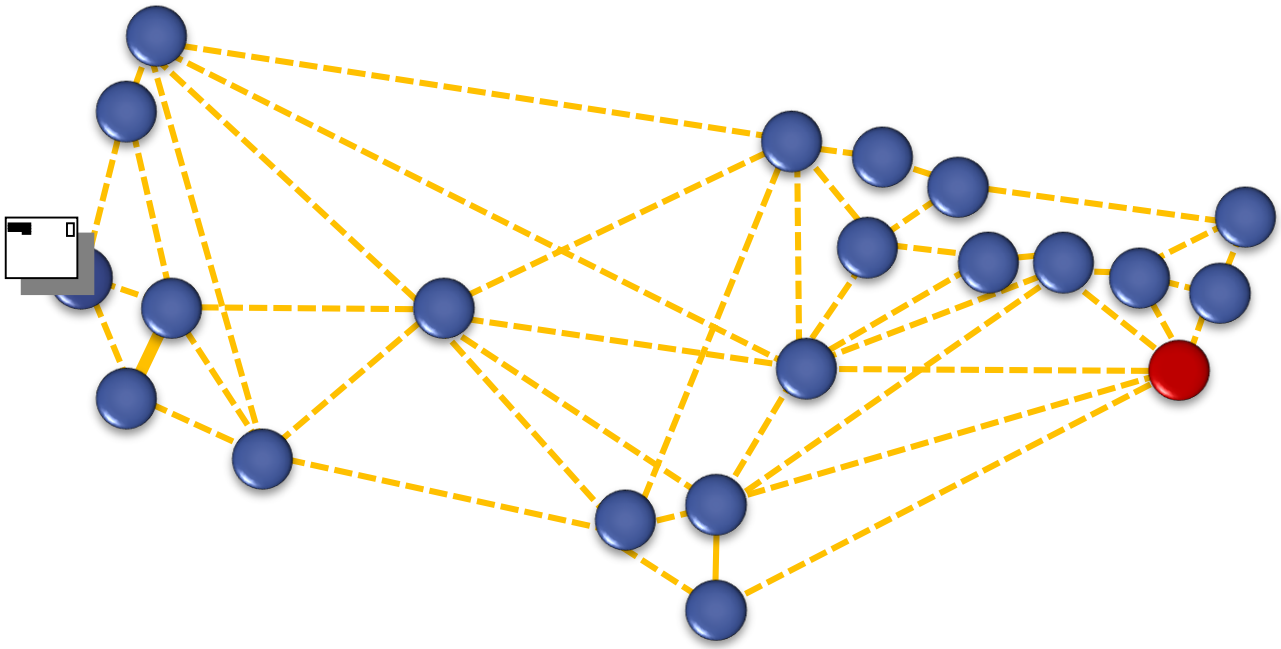
Sending the 'information' to the most connected nodes (hubs) of the network will increase the chances of reaching the target.

(see Chapter 19)

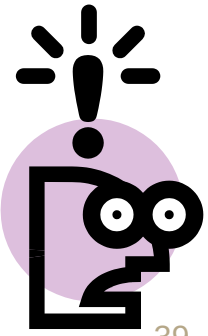
Shortest Path Approaches



But also, there is a high probability that the most connected nodes of the network are involved in a high number of transitive relations.



Sending the 'information' to the most connected nodes (hubs) of the network will increase the chances of getting lost.

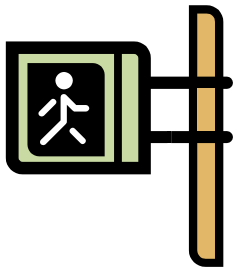


(see Chapter 19)

All-Routes Approaches

Hypothesis: The information not only flows through the shortest paths but by mean of any available route.

Definition: The communicability between two nodes in a network is defined as a function of the total *number of routes* connecting them, giving more importance to the shorter than to the longer ones.



Definition: A route is a walk of length l, which 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_l to v_{l+1}.

(see Chapter 19)

All-Routes Approaches

Communicability

The communicability function between the nodes p and q is then mathematically defined as

$$G_{pq} = \sum_{l=0}^{\infty} c_l \left(\#\ \text{of walks in}\ l\ \text{steps from}\ p\ \text{to}\ q \right)$$

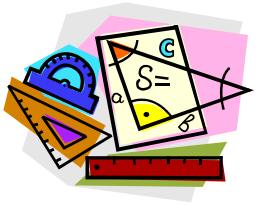
where c_l should:

- makes the series convergent
- gives more weight to the shorter than to the longer walks

(see Chapter 19)

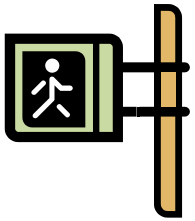
All-Routes Approaches

Communicability



Definition: Let $\ell_2(V) = \left\{ \{f(p)\}_{p\in V} \middle| \sum_{p\in V} |f(p)|^2 < \infty \right\}$.

The adjacency operator is a bounded operator on $\ell_2(V)$ defined as $(Af)(p) = \sum_{(p,q)\in E} f(q)$.

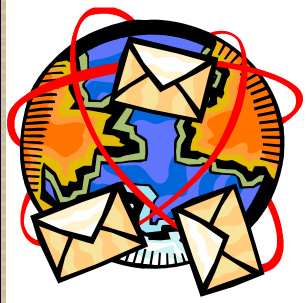


Theorem: The number of walks of length *l* between the nodes *p* and *q* in a network is equal to $\left(A^l\right)_{pq}$.

All-Routes Approaches

Communicability

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ be a nonincreasing ordering of the eigenvalues of *A*, and let $\vec{\varphi}_j$ be the eigenvector associated with λ_j .



The *communicability function* can be written as:

$$G_{pq} = \sum_{l=0}^{\infty} c_l \left(A^l\right)_{pq} = \sum_{l=0}^{\infty} \sum_{j=1}^{n} c_l \lambda_j^l \varphi_{j,p} \varphi_{j,q}$$

(see Chapters 12 & 19)

All-Routes Approaches

Communicability

Factorial penalisation

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

Parametric penalisation

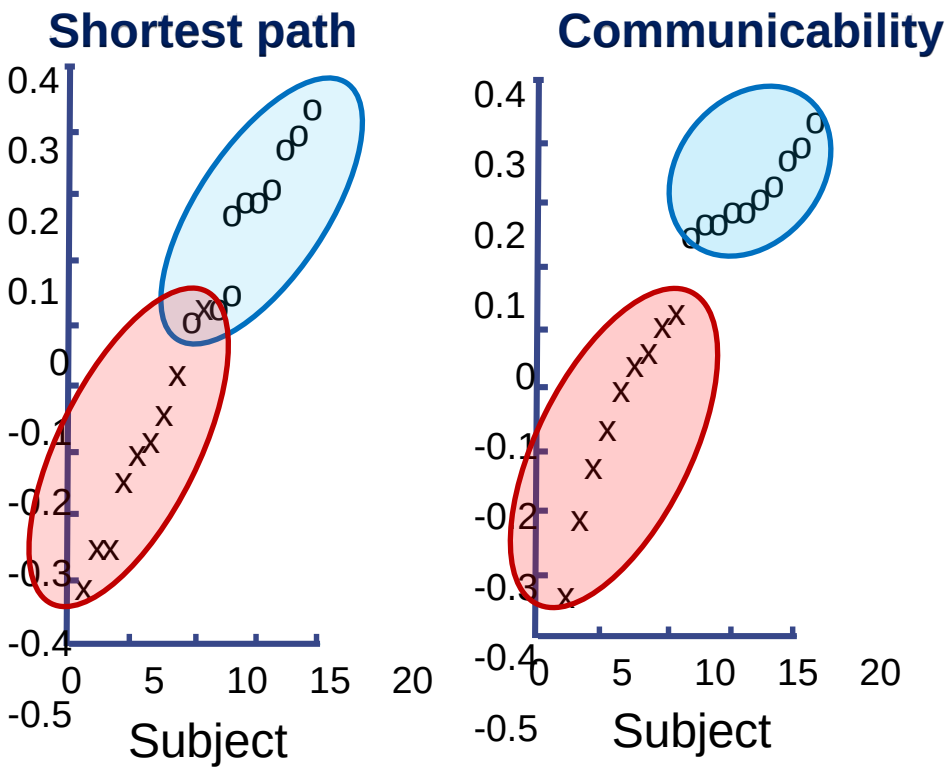
$$G_{pq} = \sum_{l=0}^{\infty} \alpha^l \left(A^l\right)_{pq} = \left(\left(I - \alpha A\right)^{-1}\right)_{pq} = \sum_{j=1}^{n} \frac{\varphi_{j,p}\varphi_{j,q}}{1 - \alpha\lambda_j}$$

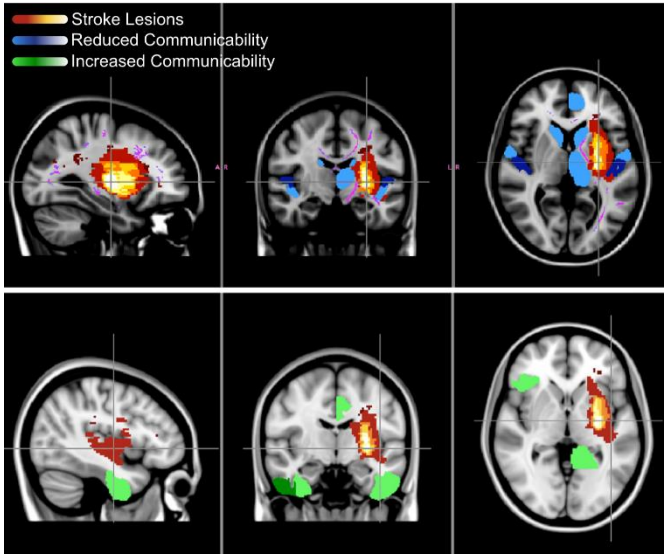
$$0 < \alpha < \lambda_1^{-1}$$

(see Chapters 12 & 19)

Experimental evidences

Stroke lesions





Crofts & Higham: *J. Roy. Soc. Interface* **6** (2009) 411
Crofts et al: *NeuroImage* **54** (2011) 161-9

Experimental evidences

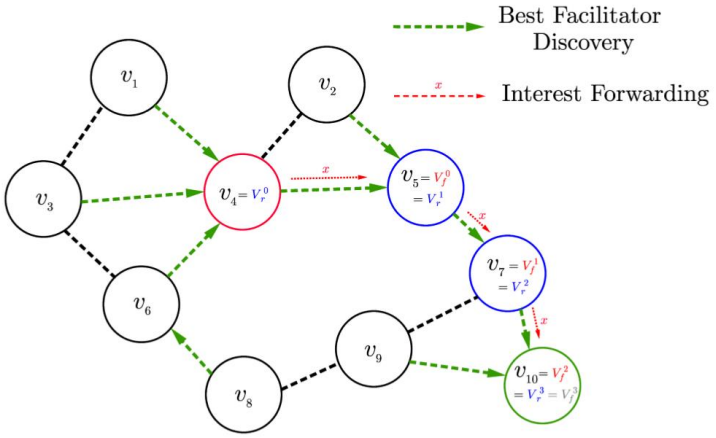


Figure 1: Facilitator Discovery Process

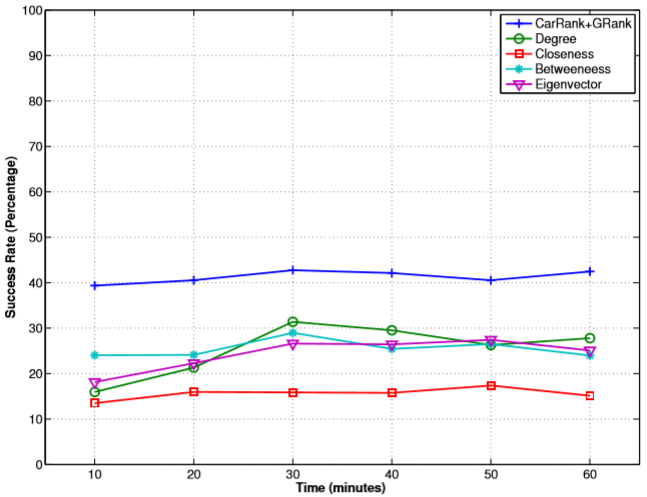
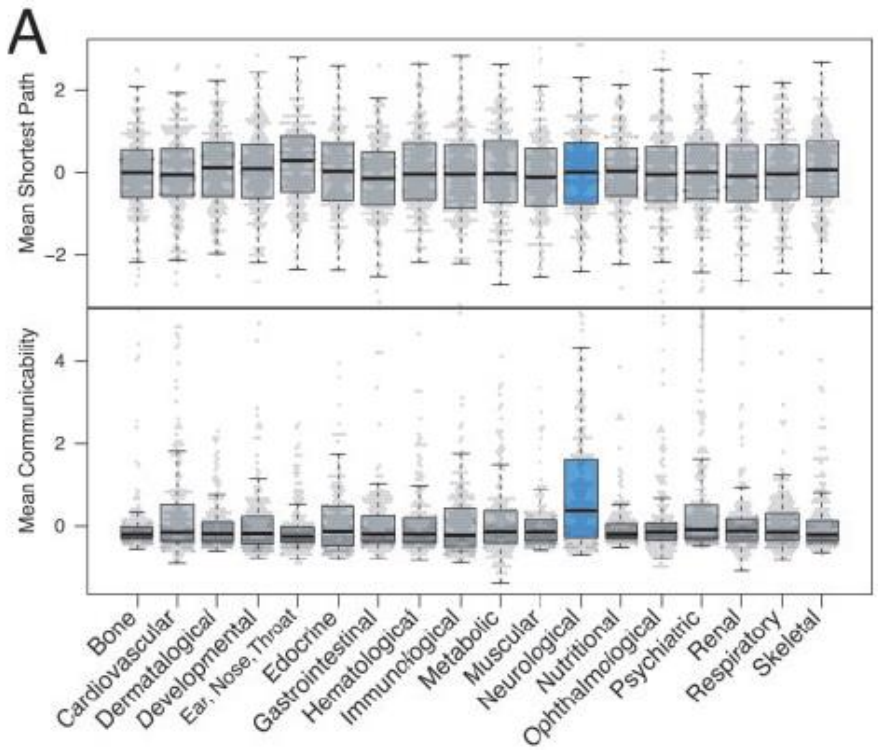


Figure 3: Success rate conparison for the consumer interests satisfied over time using different centrality schemes for content distribution

Inspired from the concept of communicability in complex networks, Grank, a global vehicle scheme allows a vehicles to use a new stable metric named "Information communicability" to rank different locations in the city and rank itself accordingly.

Khan, and Ghamri-Doudane. *IEEE Comm. Mag.* 54.8 (2016): 100-107.

Experimental evidences

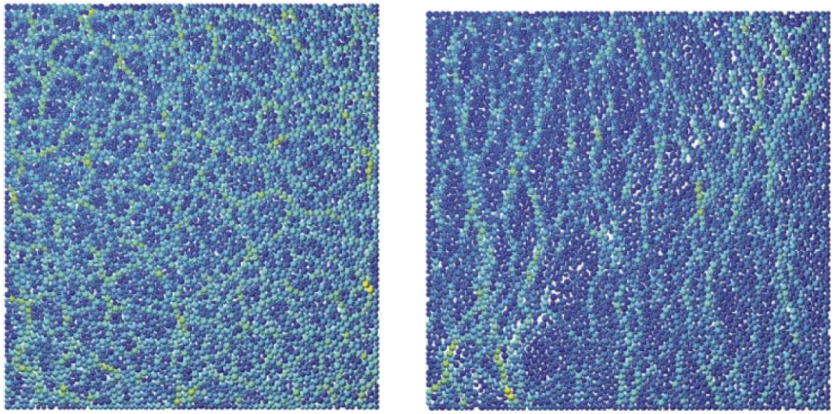


Network communicability provides advantages over alternative metrics because it retains topology information, lends itself to set-based analysis and it is easy to represent with univariate scores. It outperforms shortest path in a variety of situations.

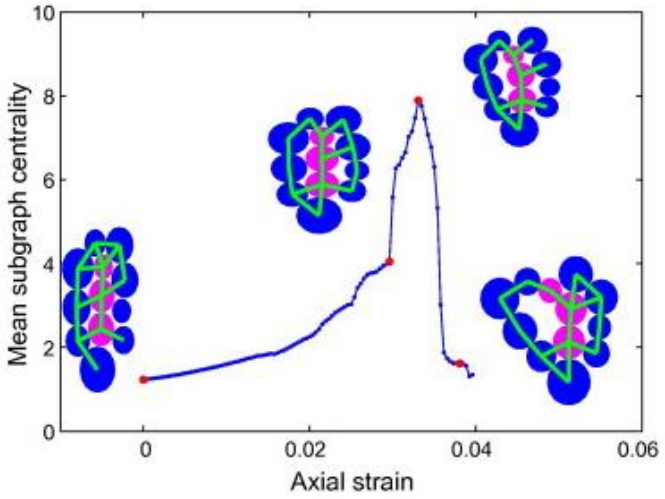
Campbell et al: *Bioinformatics* **30** (2014) 3387-3389

Experimental evidences

Granular materials



Effects of axial strain



(…) weighted subgraph centrality measure correlates strongly with nonaffine deformation and dissipation, spatially and temporally, and at both the mesoscopic and macroscopic level. (…) a large decrease in the value of the weighted subgraph centrality occurs when a force chain buckles and its local neighboring particles rearrange.

I'll be back!

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

LECTURE 3: BASED ON CHAPTERS 11 & 12

Clustering coefficients

Ernesto Estrada

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

Recall

Steven Strogatz

Definition 1. *Let* $G=(V,E)$ *be a connected, simple, undirected network, and let* $i \in V$. *Then, the clustering coefficient of the node* i *is given by*

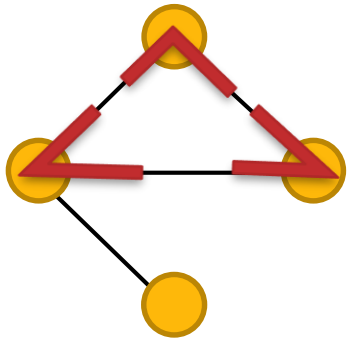
$$C_i = \frac{t_i}{k_i(k_i-1)/2} = \frac{2t_i}{k_i(k_i-1)}.$$

Duncan Watts

The average WS clustering is then given by,

$$\overline{C} = \frac{1}{n}\sum_i C_i$$

Transitivity index



Each triangle contains three paths of length two

Definition 2. *Let* $G = (V, E)$ *be a connected, simple, undirected network. Let* $|C_3|$ *and* $|P_2|$ *be the total number of triangles and of paths of length two in* G*. Then, the global clustering* $C = C(G)$ *of* G *is given by*

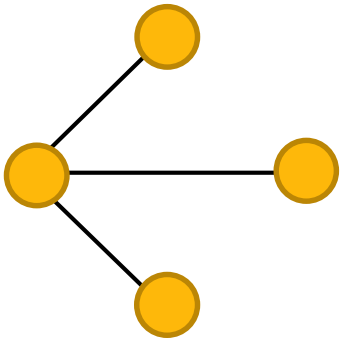
$$C = \frac{3|C_3|}{|P_2|}.$$

Transitivity index

Recall that:

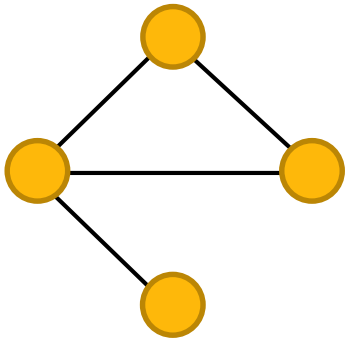
$$|C_3| = \frac{1}{6} tr\left(A^3\right) \qquad |P_2| = \frac{1}{2} \sum_{i=1}^{n} k_i \left(k_i - 1\right)$$

Example 3



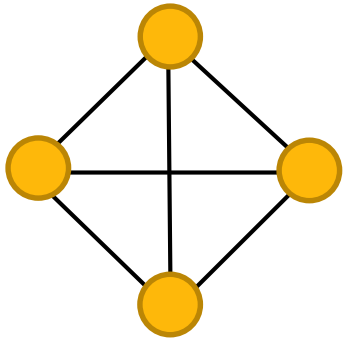
$C = 0$

$\overline{C} = 0$



$C = 1/5$

$\overline{C} = 7/12$

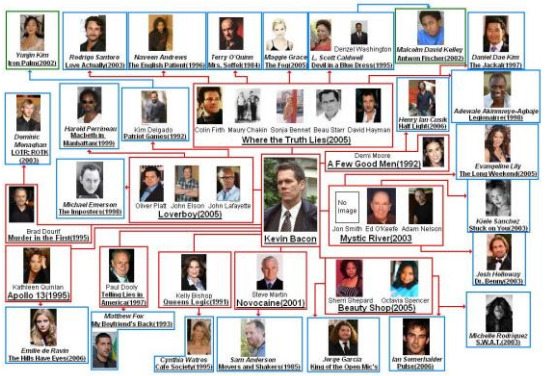


$C = 1$

$\overline{C} = 1$

Average WS vs. transitivity

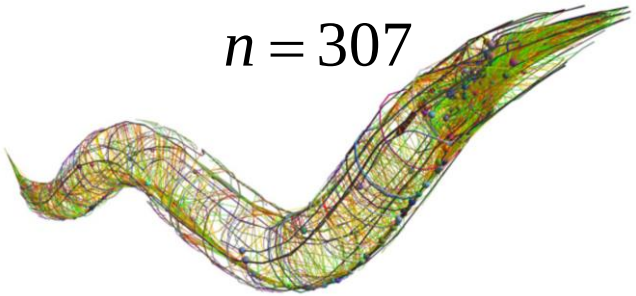
$n = 449{,}913$



$\overline{C} = 0.78$

$C = 0.20$

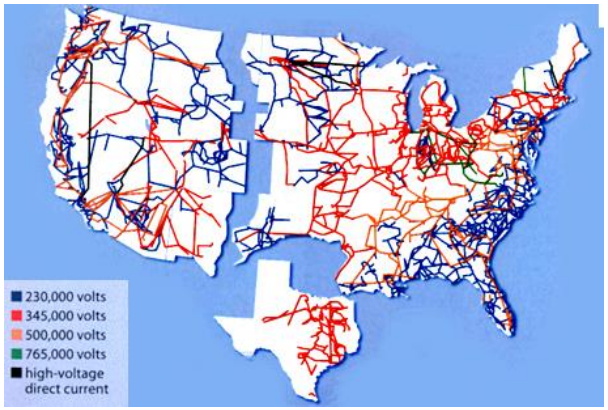
$n = 307$



$\overline{C} = 0.28$

$C = 0.18$

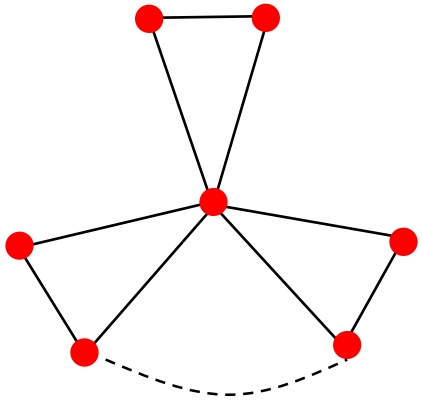
$n = 4{,}941$



$\overline{C} = 0.08$

$C = 0.10$

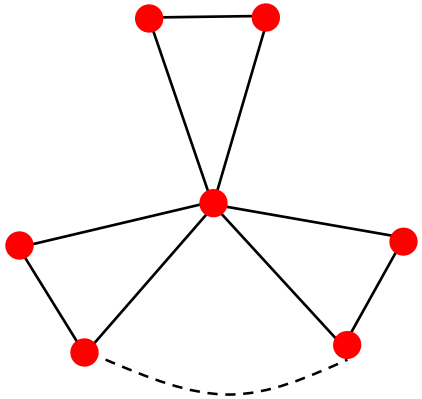
Average WS vs. transitivity



Lemma 1. *Let* $G = (V, E)$ *be a friendship graph. Then, the average Watt-Strogatz and the global clustering coefficients are given, respectively, by*

$$\overline{C} = \frac{1}{n(n-2)} + \frac{n-1}{n} \qquad C = \frac{3}{n}$$

Is there a Problem?



Let

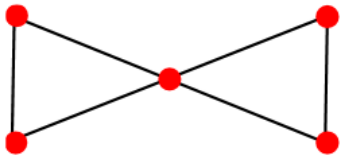
$$\lim_{n\to\infty}\overline{C}=\lim_{n\to\infty}\left(\frac{1}{n(n-2)}+\frac{n-1}{n}\right)=1$$

$$\lim_{n\to\infty}C=\lim_{n\to\infty}\frac{3}{n}=0$$

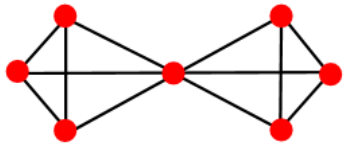
Is that a specific disease?

Definition 4. *A windmill graph* $W(\eta, \kappa)$ *of* $n = \eta\kappa + 1$ *nodes consists of* η *copies of the complete graph* K_κ *with every node connected to a node* $i \notin K_\kappa$.

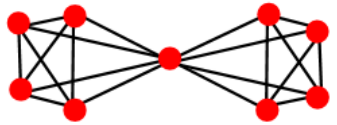
Example 4



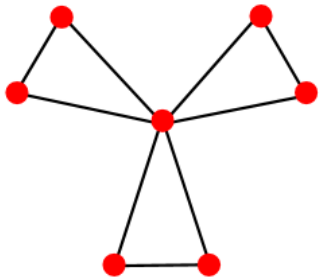
$W(2,2)$



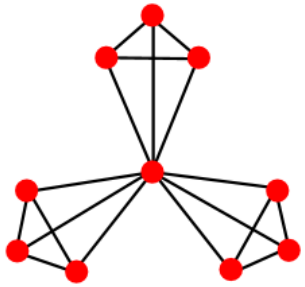
$W(2,3)$



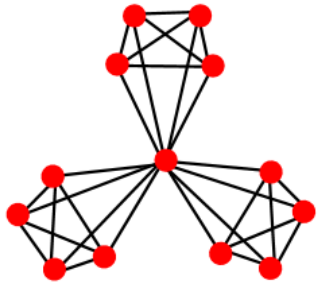
$W(2,4)$



$W(3,2)$



$W(3,3)$



$W(3,4)$

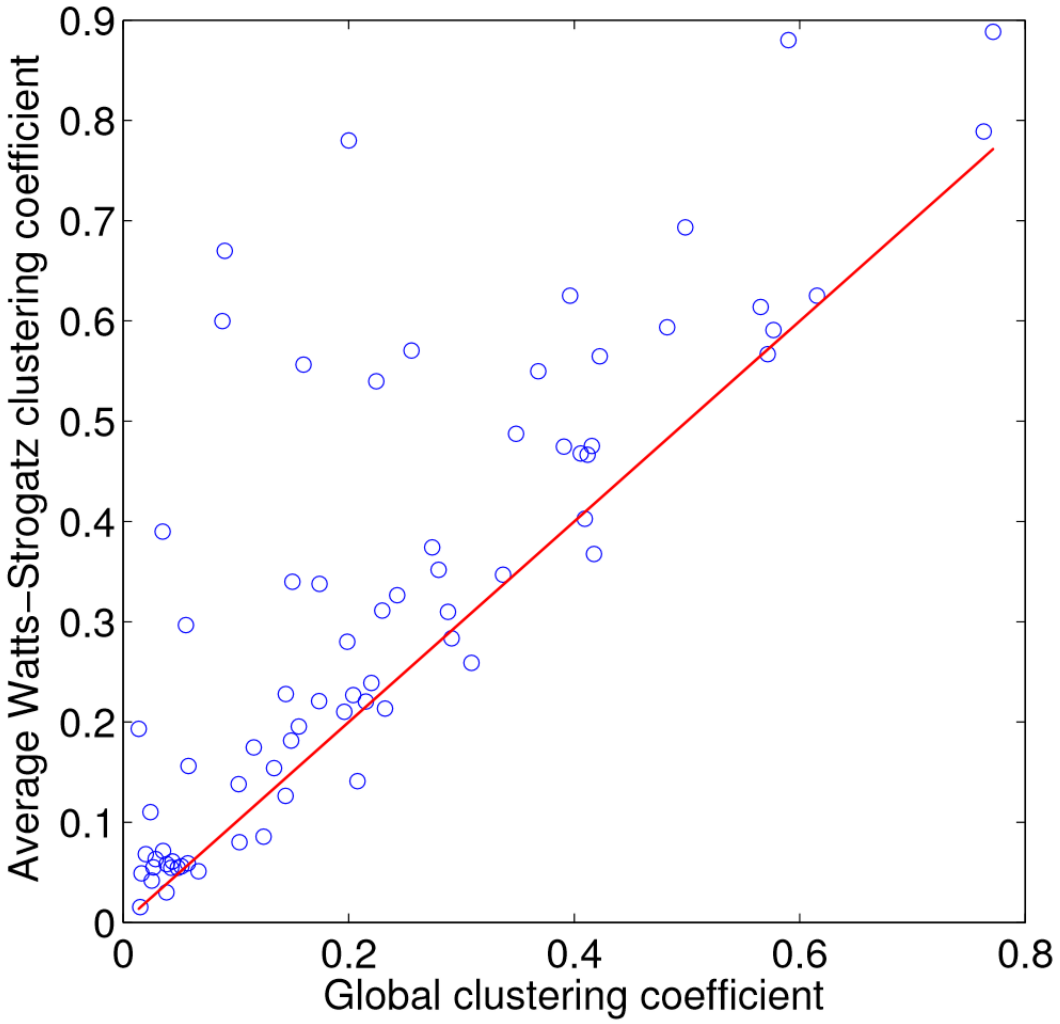
Is that a specific 'disease'?

Theorem 3. *Let* $W = W(\eta, \kappa)$ *be a windmill graph. Then, for a given value of* κ*, the average Watts-Strogatz and the transitivity indices diverge when the number of cliques tends to infinity,*

$$\lim_{\eta \to \infty} \overline{C}(W) = 1 \qquad \lim_{\eta \to \infty} C(W) = 0$$

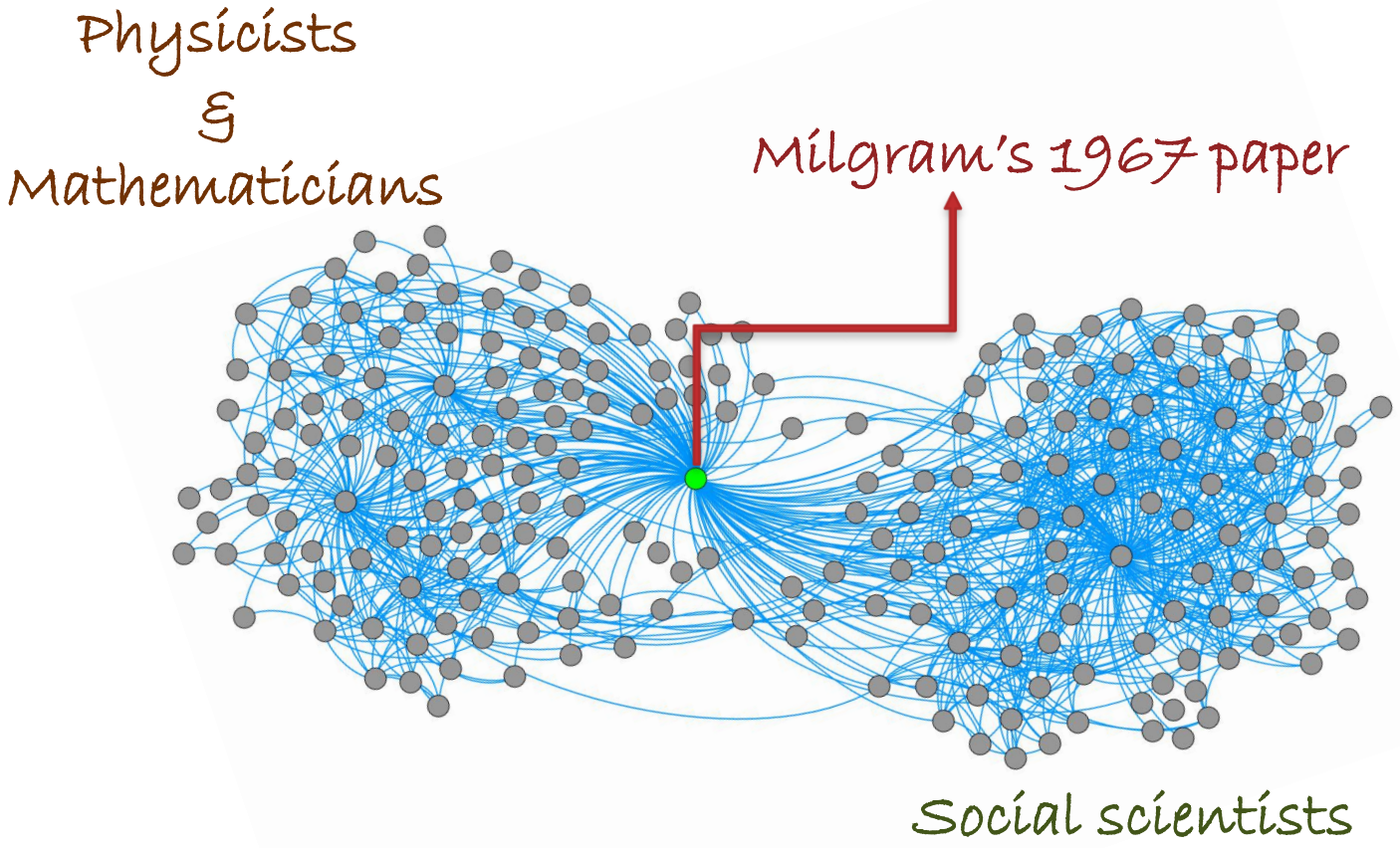
Is the real-world affected?

Real-world networks



Is the real-world affected?

Citation network of "Small-World"

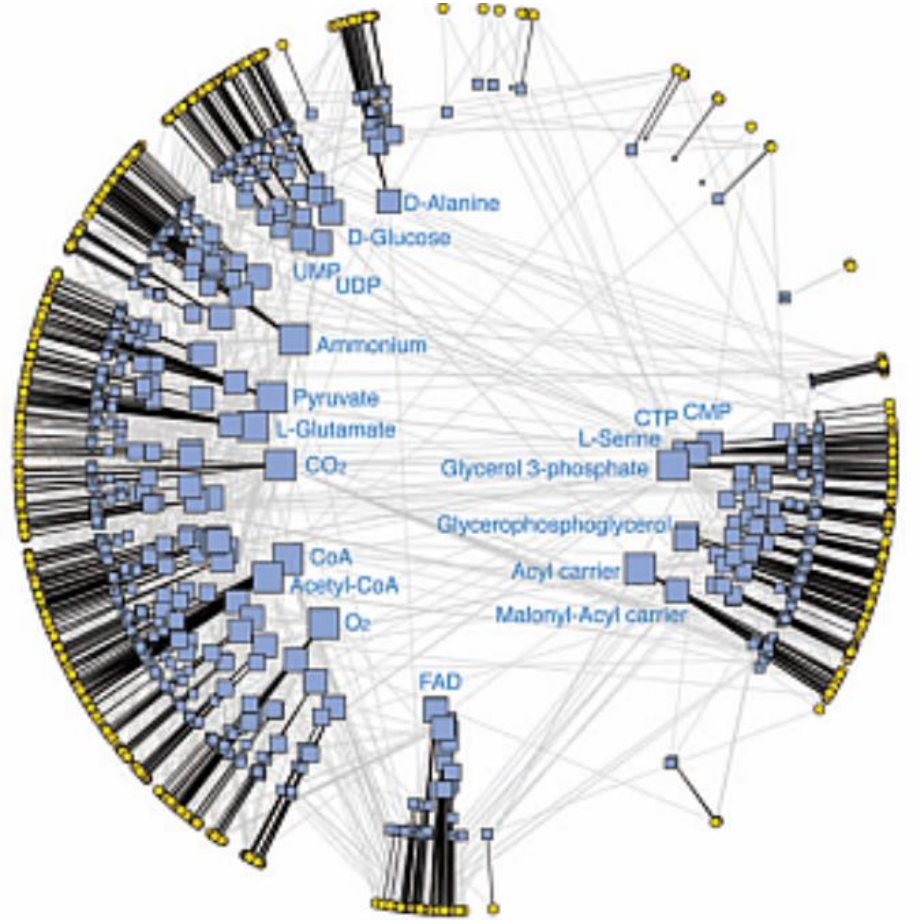


$$\overline{C} = 0.56$$

$$C = 0.16$$

Is the real-world affected?

Metabolic network of *E. coli*



$$\overline{C} = 0.60$$

$$C = 0.09$$

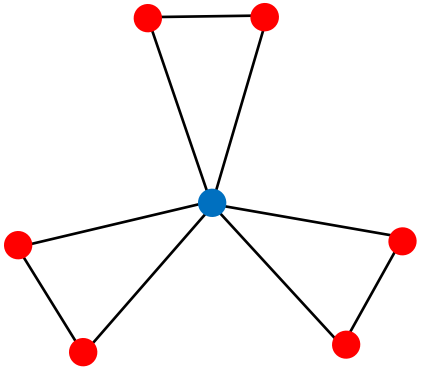
Can we recover sanity?

Remark 1. *The average WS clustering coefficient measures how locally 'clustered' a network is on average.*

Remark 2. *The transitivity index measures how globally 'clustered' a network is.*

Can we recover sanity?

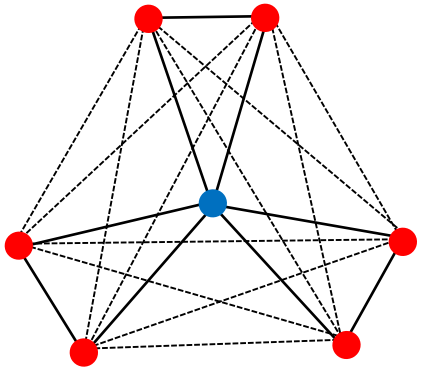
Example 6



$C_{\bullet} = 1$

$C_{\bullet} = 1/5$

$\overline{C} \approx 0.74$



$C = 3/7 \approx 0.42$

LECTURE 4: BASED ON CHAPTERS 11 & 19

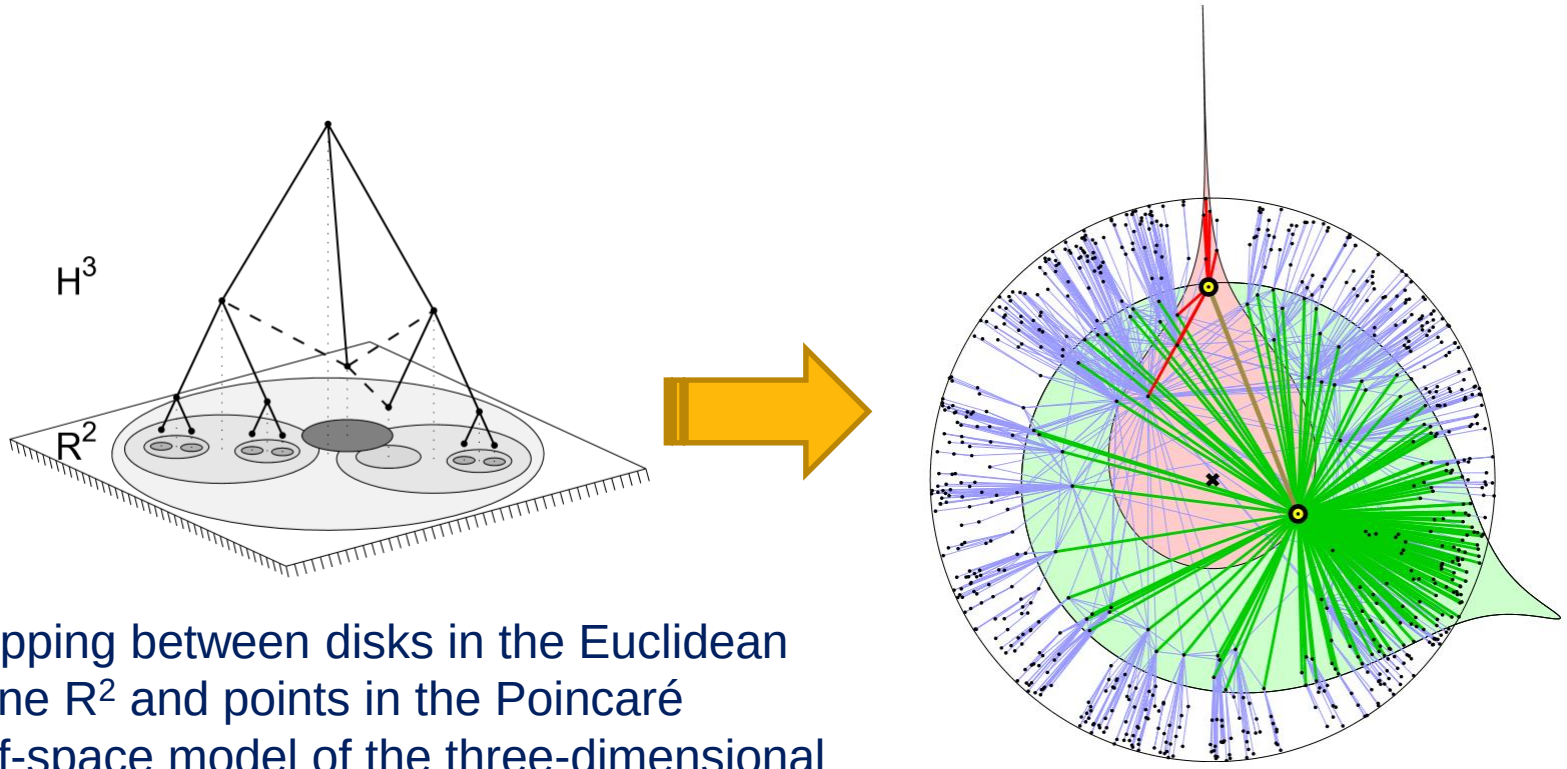
Imposed vs. Emerged Geometry

Ernesto Estrada

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

Imposed embedding



Mapping between disks in the Euclidean plane R^2 and points in the Poincaré half-space model of the three-dimensional hyperbolic space H^3

Krioukov, et al. *Phys. Rev. E* 82 (2010): 036106.

University of Strathclyde

Imposed embedding

Hyperbolic Geometry of Complex Networks

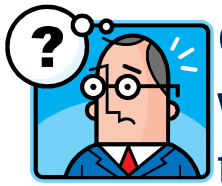
Dmitri Krioukov,[1] Fragkiskos Papadopoulos,[2] Maksim Kitsak,[1] Amin Vahdat,[3] and Marián Boguñá[4]

[1] *Cooperative Association for Internet Data Analysis (CAIDA), University of California, San Diego (UCSD), La Jolla, CA 92093, USA*
[2] *Department of Electrical and Computer Engineering, University of Cyprus, Kallipoleos 75, Nicosia 1678, Cyprus*
[3] *Department of Computer Science and Engineering, University of California, San Diego (UCSD), La Jolla, CA 92093, USA*
[4] *Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*

We develop a geometric framework to study the structure and function of complex networks. We assume that hyperbolic geometry underlies these networks, and we show that with this assumption, heterogeneous degree distributions and strong clustering in complex networks emerge naturally as simple reflections of the negative curvature and metric property of the underlying hyperbolic geometry. Conversely, we show that if a network has some metric structure, and if the network degree distribution is heterogeneous, then the network has an effective hyperbolic geometry underneath. We then establish a mapping between our geometric framework and statistical mechanics of complex networks. This mapping interprets edges in a network as non-interacting fermions whose energies are hyperbolic distances between nodes, while the auxiliary fields coupled to edges are linear functions of these energies or distances. The geometric network ensemble subsumes the standard configuration model and classical random graphs as two limiting cases with degenerate geometric structures. Finally, we show that targeted transport processes without global topology knowledge, made possible by our geometric framework, are maximally efficient, according to all efficiency measures, in networks with strongest heterogeneity and clustering, and that this efficiency is remarkably robust with respect to even catastrophic disturbances and damages to the network structure.

Why a mapping into H^3?

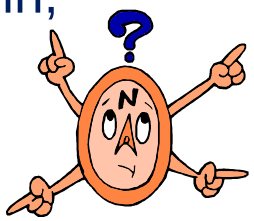
A more natural approach...



Consider that two nodes p and q are trying to communicate with each other by sending information both ways through the network.

We know that:

G_{pp} Quantifies the amount of 'information' that is sent by p, wanders around the network, and returns to the origin, the node p.

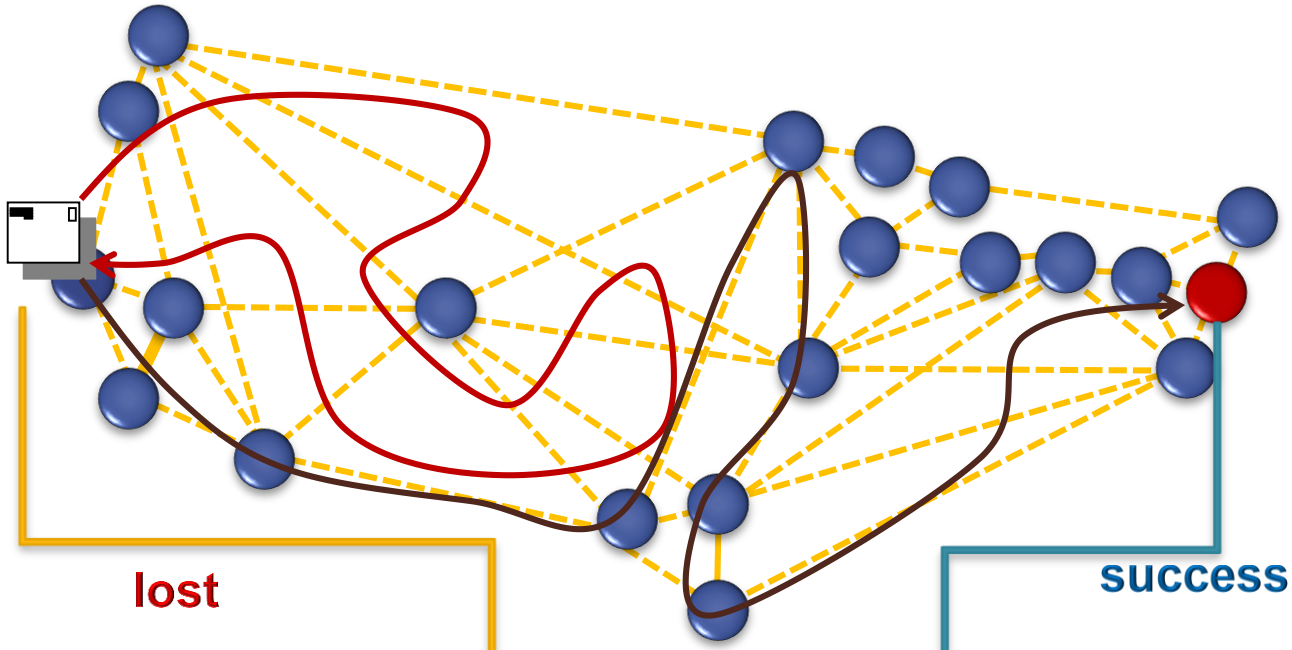


G_{pq} Quantifies the amount of 'information' that is sent by p, wanders around the network, and arrives to its destination, the node q.

(see Chapter 19)

A more natural approach...

The goal of communication is to maximise the amount of information that arrives to its destination by minimising that which is lost.



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

(see Chapter 19)

Communicability Distance

Theorem: The function ξ_{pq} is a Euclidean distance that induces a natural embedding of the network into a Euclidean space.

Proof:

$$\xi_{pq} = \left(\boldsymbol{\varphi}_p - \boldsymbol{\varphi}_q\right)^T e^{\boldsymbol{\Lambda}} \left(\boldsymbol{\varphi}_p - \boldsymbol{\varphi}_q\right)$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \qquad \boldsymbol{\varphi}_p = \begin{bmatrix} \varphi_1(p) \\ \varphi_2(p) \\ \vdots \\ \varphi_n(p) \end{bmatrix}$$

(see Chapter 19)

Communicability Distance

Proof (cont.):

$$\xi_{pq} = \left[e^{\mathbf{\Lambda}/2} \left(\boldsymbol{\varphi}_p - \boldsymbol{\varphi}_q \right) \right]^T e^{\mathbf{\Lambda}/2} \left(\boldsymbol{\varphi}_p - \boldsymbol{\varphi}_q \right)$$

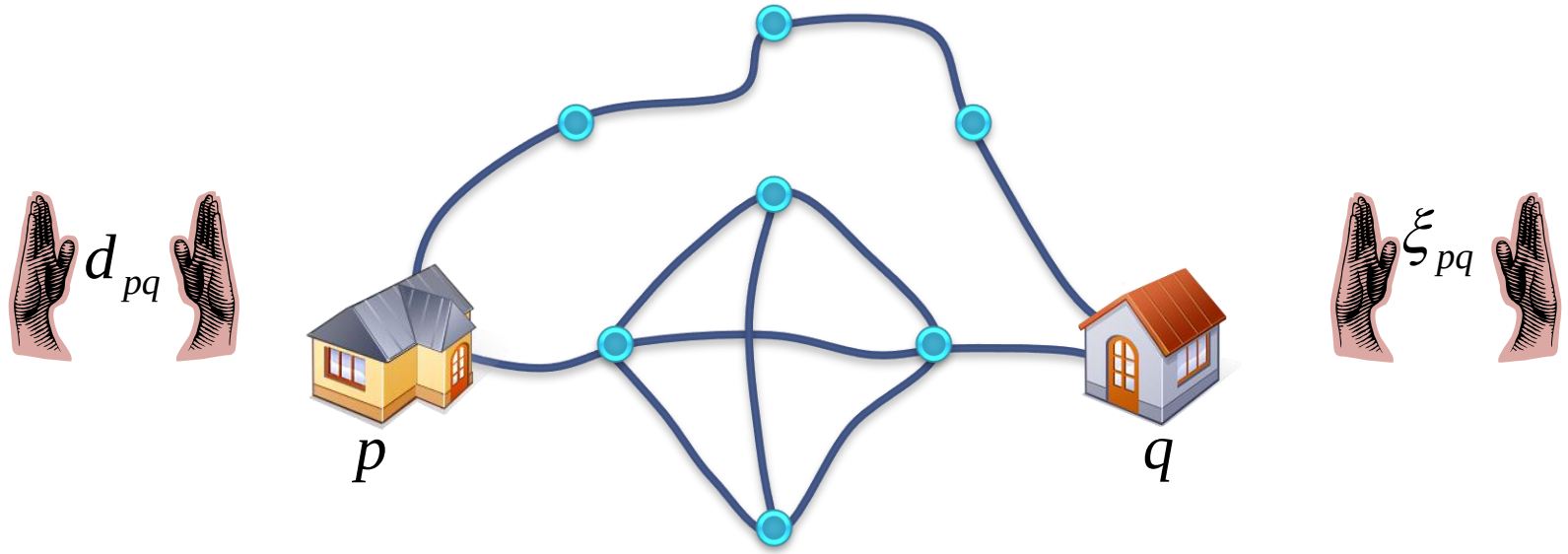
$$= \left(e^{\mathbf{\Lambda}/2} \boldsymbol{\varphi}_p - e^{\mathbf{\Lambda}/2} \boldsymbol{\varphi}_q \right)^T \left(e^{\mathbf{\Lambda}/2} \boldsymbol{\varphi}_p - e^{\mathbf{\Lambda}/2} \boldsymbol{\varphi}_q \right)$$

$$= \left(\mathbf{x}_p - \mathbf{x}_q \right)^T \left(\mathbf{x}_p - \mathbf{x}_q \right)$$

$$= \left\| \mathbf{x}_p - \mathbf{x}_q \right\|^2$$

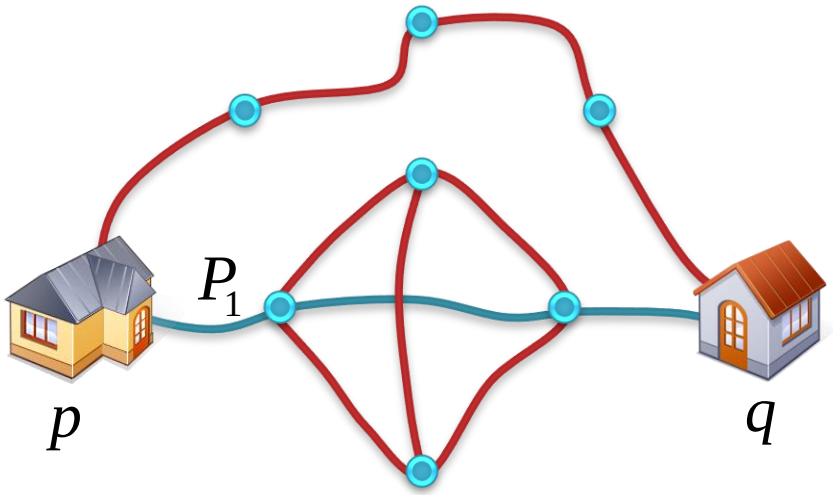
(see Chapter 19)

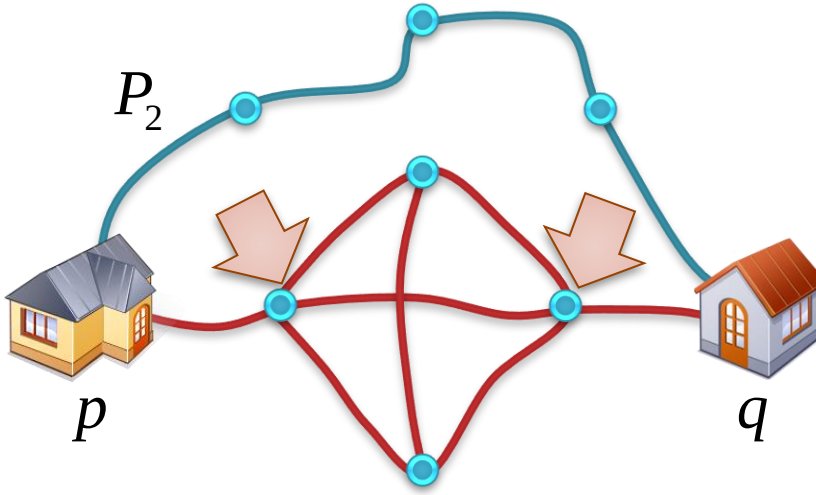
Communicability Distance



(see Chapter 19)

Communicability Paths





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

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

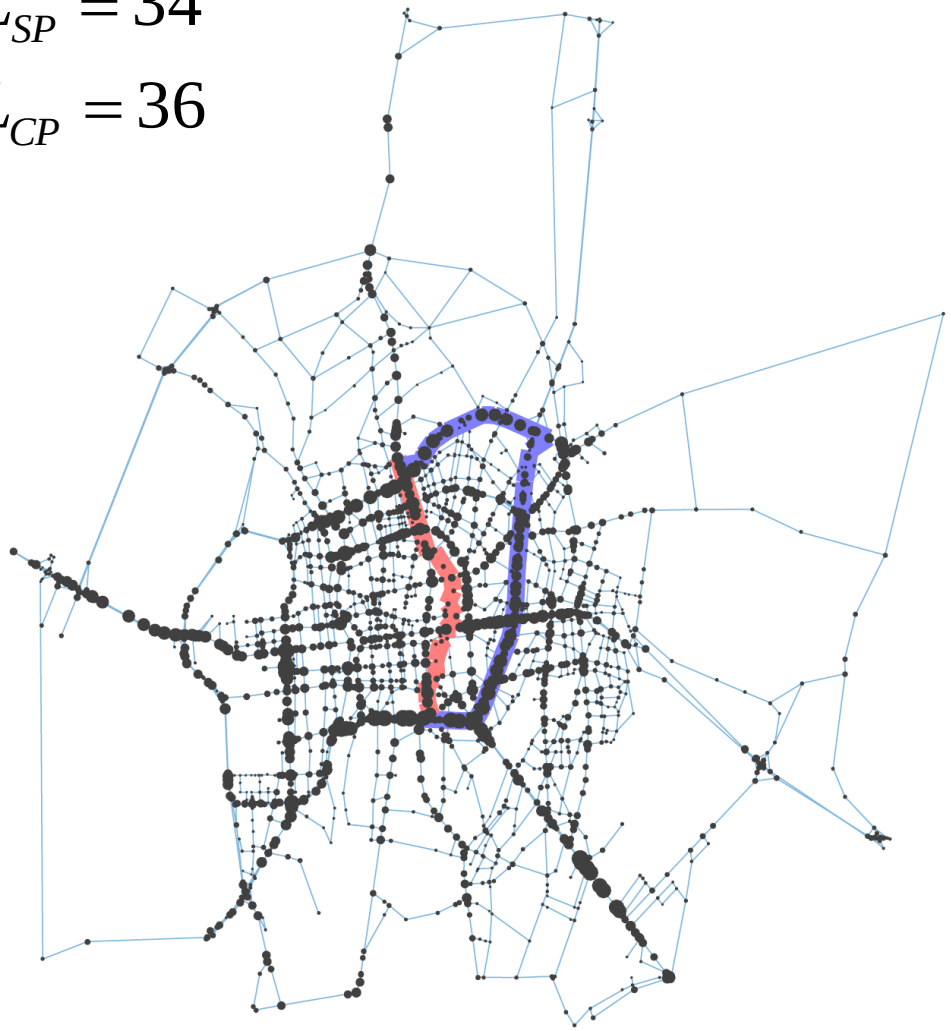
Remark. *The shortest route based on the communicability distance is the shortest path that avoids the nodes with the highest 'cliquishness' in the graph.*

(see Chapter 19)

Communicability geometry

$L_{SP} = 34$

$L_{CP} = 36$



Isfahan, Iran
1,961,260 inhabitants

Shortest path/Comm path

$r_{BC} = 1.00$

$r_{CC} = 1.07$

$r_{SC} = 1.12$

$r_{DC} = 1.69$

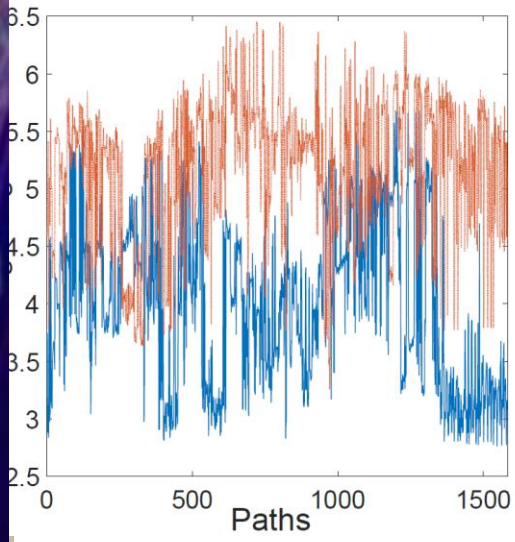
$r_{EC} = 7.03$

$$r_C = \frac{C_{SP}}{C_{CP}}$$

Communicability geometry

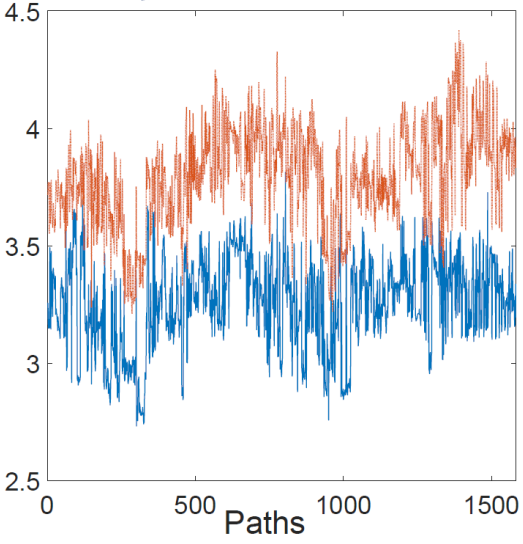
1,583 pairs of locations in the city connected by the shortest path and by the shortest communicability route

degree



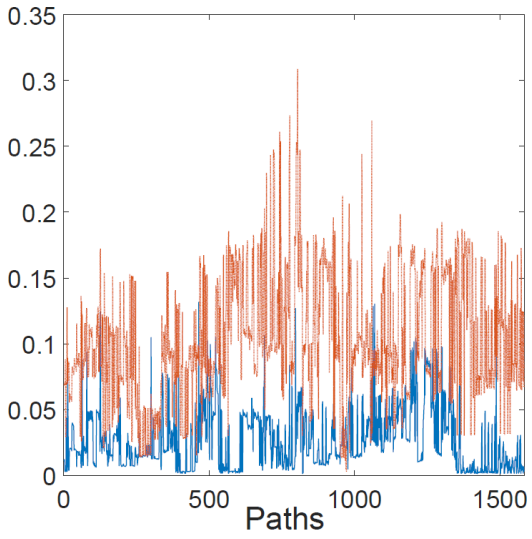
29.5% higher in SP

subgraph centrality



16.8% higher in SP

eigenvector centrality



289.6% higher in SP

$SP \ \langle L \rangle = 45$

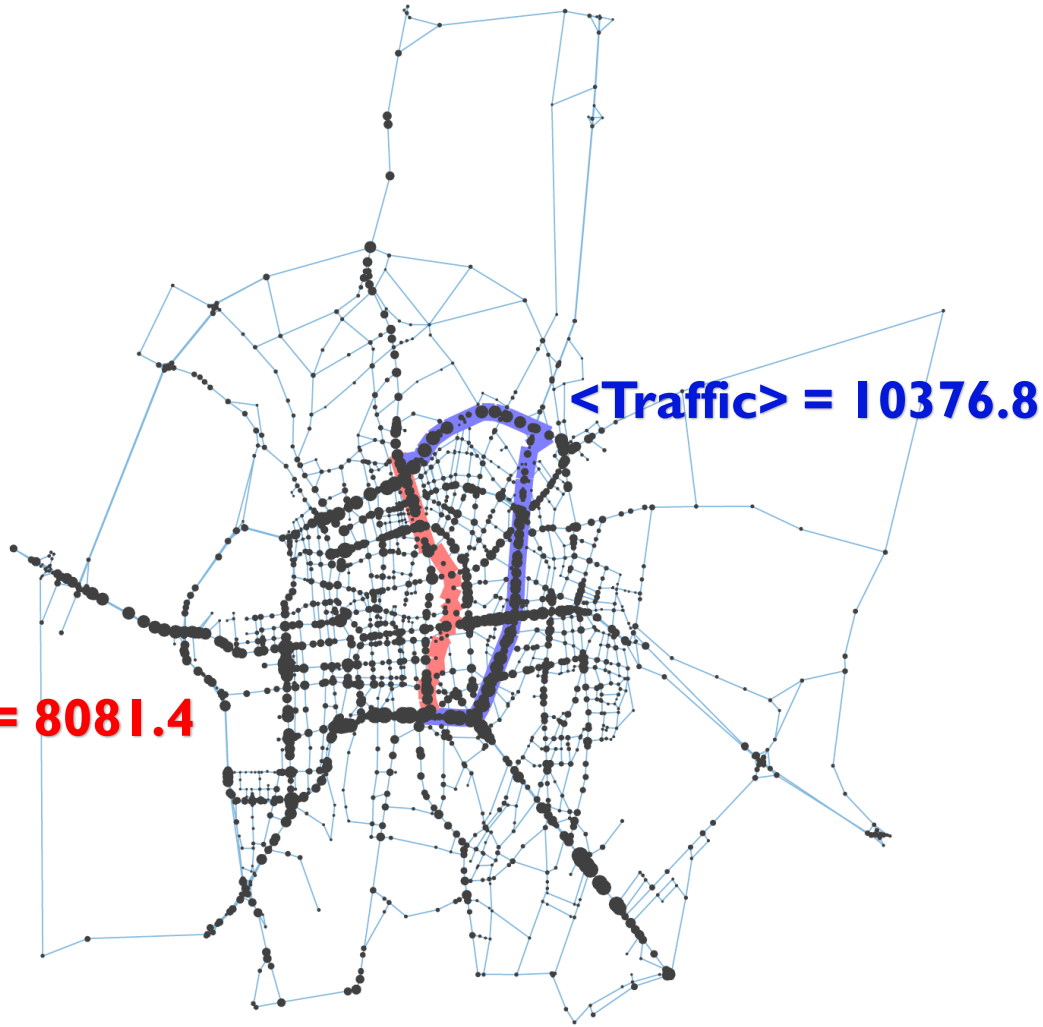
$CP \ \langle L \rangle = 47$

Communicability geometry

Number of cars per hour in the morning at intersection points

$L_{SP} = 34$

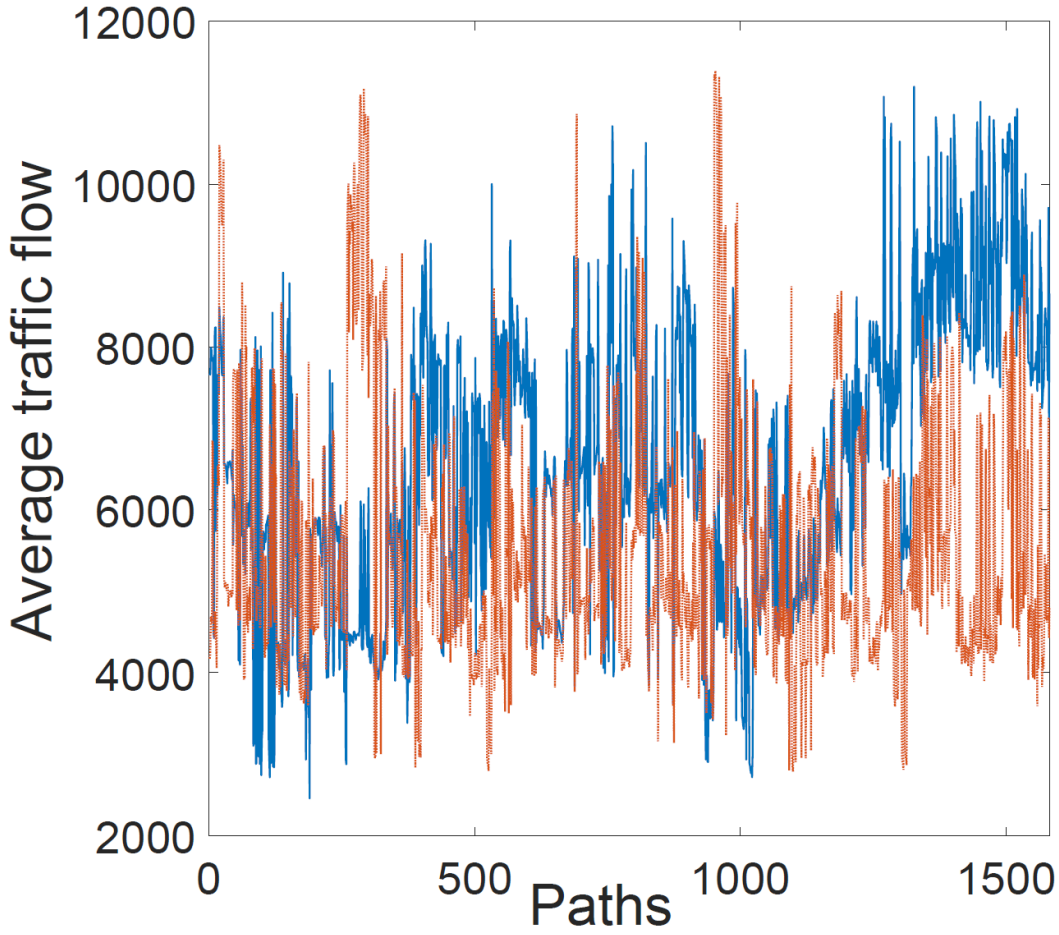
$L_{CP} = 36$



Isfahan, Iran
1,961,260 inhabitants

Communicability geometry

Isfahan, Iran
1,961,260 inhabitants

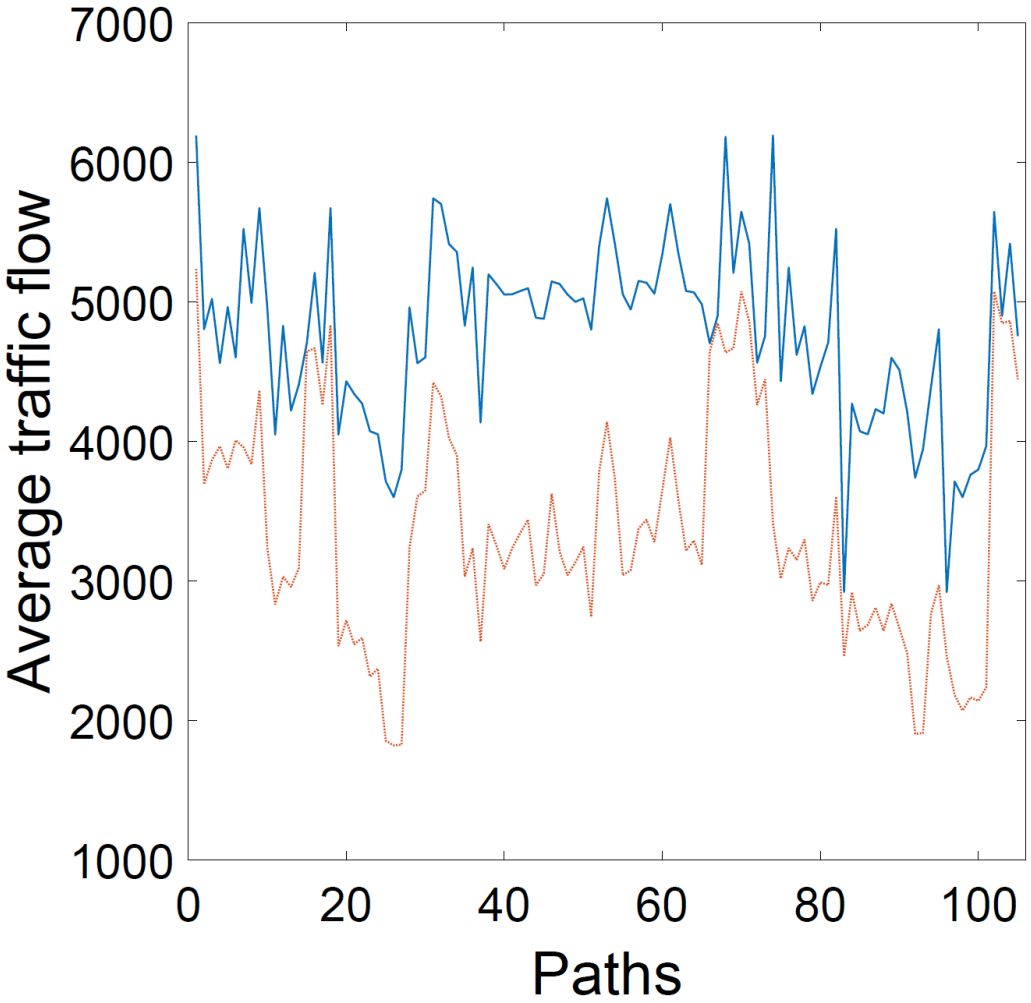


15.1% more average traffic through the communicability routes than through SP

Are people in Isfahan mad?

Communicability geometry

Anaheim, USA
336,265 inhabitants



35.2% more average traffic through the communicability routes than through SP

Remark. *People not always use the shortest path!*

Communicability angle

Definition:

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

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

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

Communicability Geometry

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

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

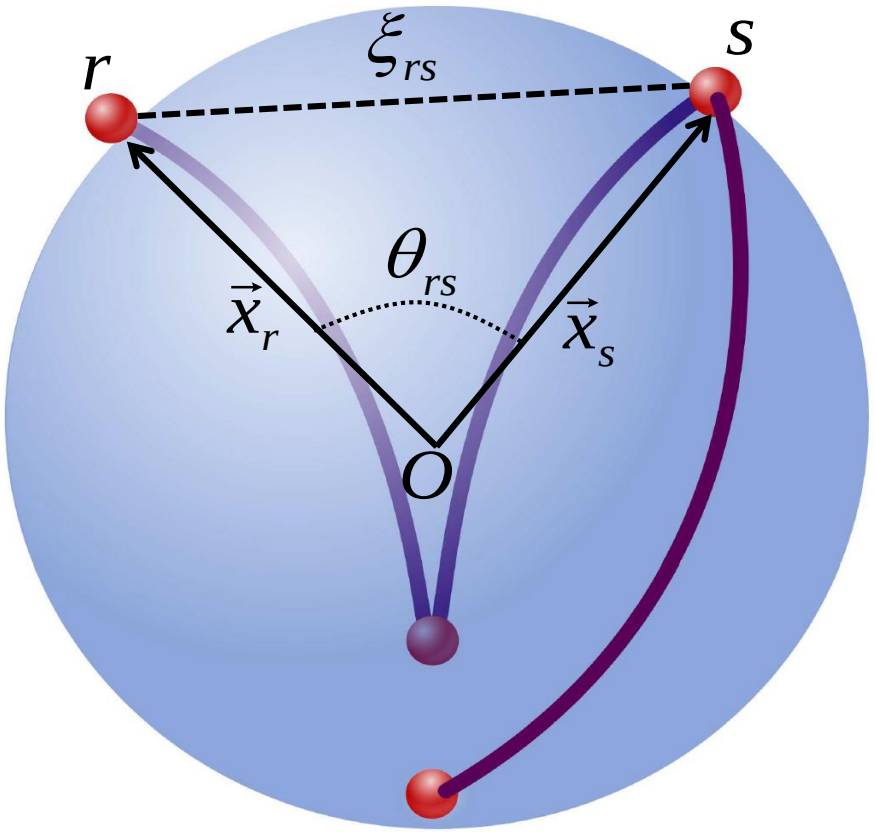
$$a = \vec{1}^T e^{-A} \vec{1} \qquad b = \vec{s}^T e^{-A} \vec{1} \qquad c = \vec{s}^T e^{-A} \vec{s}$$

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

Communicability Geometry



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

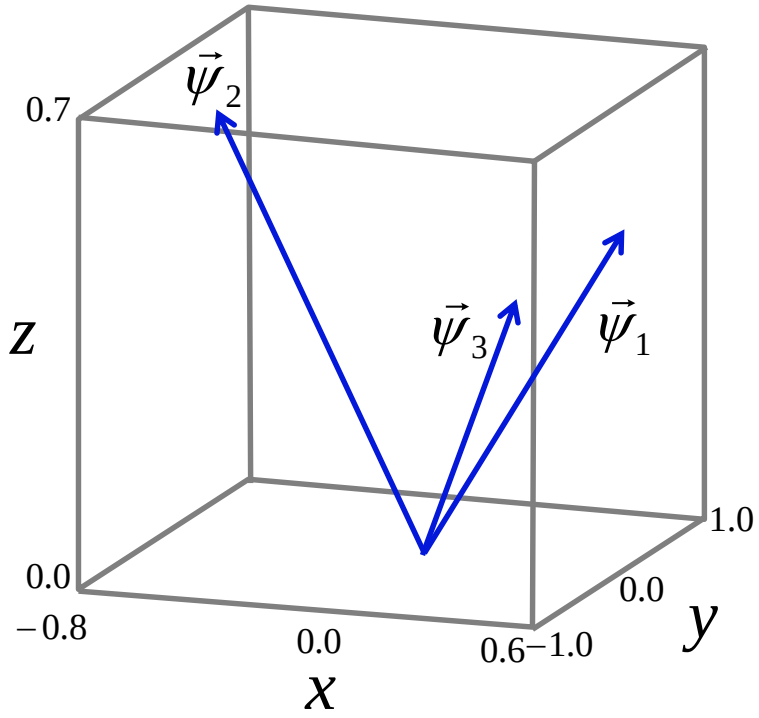
Communicability geometry

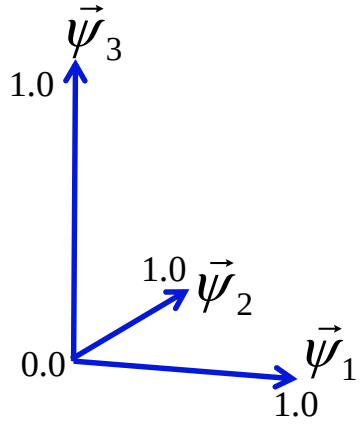
Example:

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

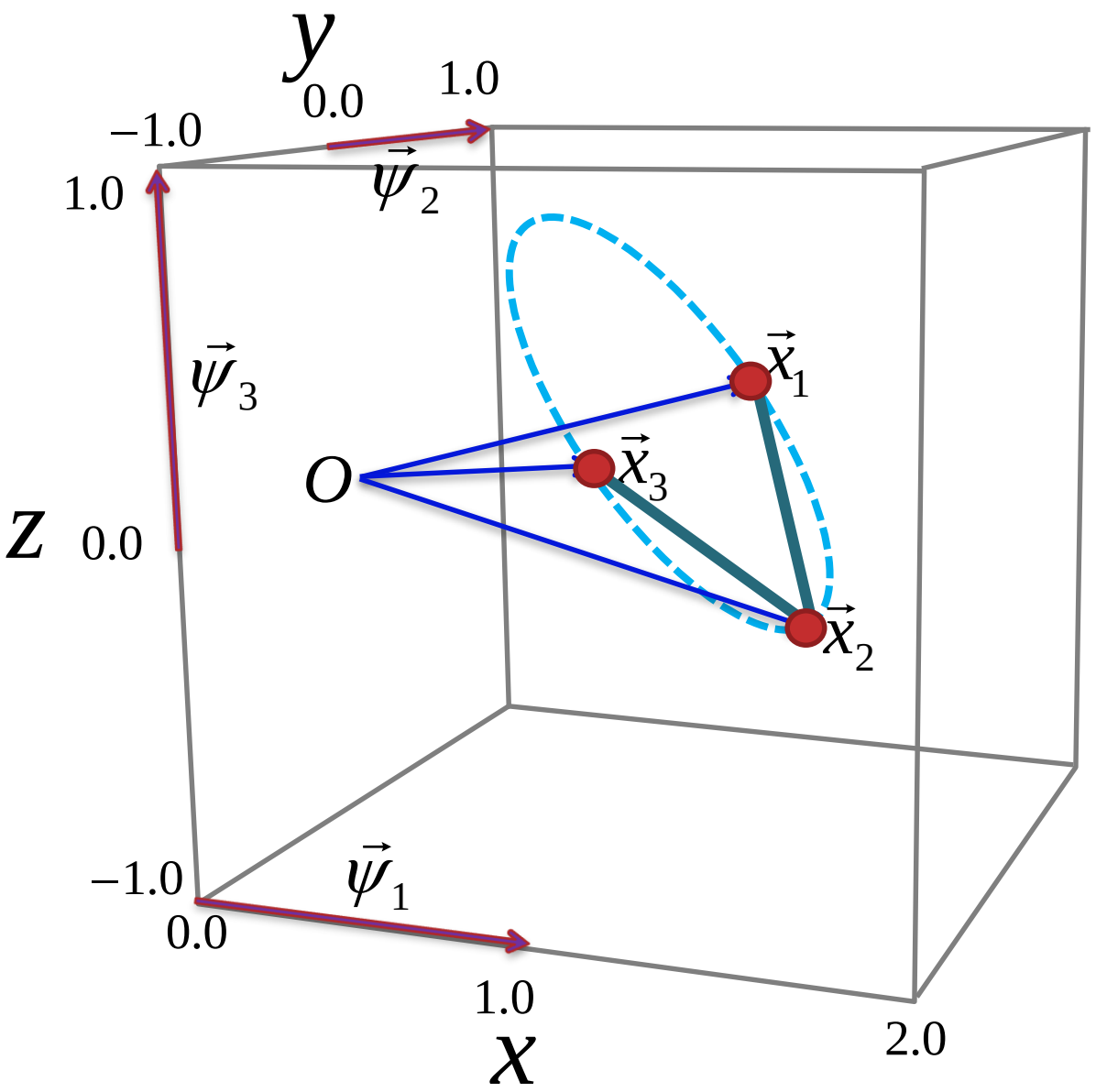
$$\vec{\psi}_1 = \begin{bmatrix} 1/2 \\ 1/\sqrt{2} \\ 1/2 \end{bmatrix} \qquad \vec{\psi}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \qquad \vec{\psi}_3 = \begin{bmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{bmatrix}$$

Communicability Geometry

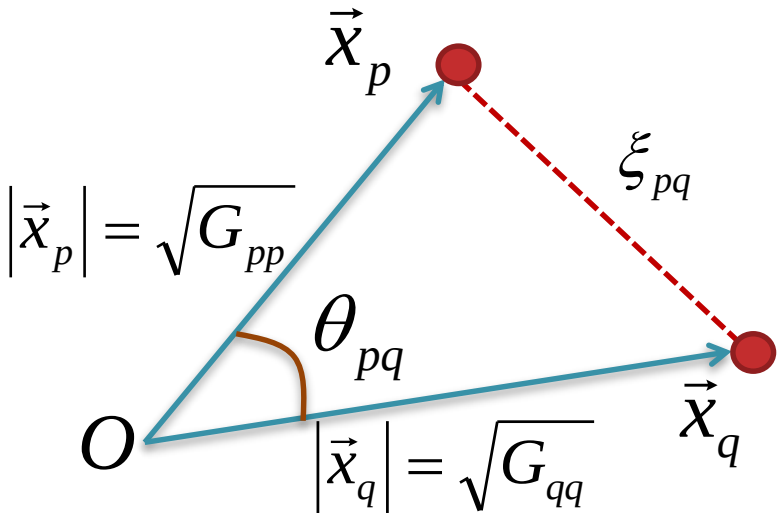




Communicability Geometry



Communicability Geometry



Communicability function

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

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

LECTURE 3: BASED ON CHAPTER 18 & 21

Network Bipartivity. Computational vs. Algebraic Approaches

Ernesto Estrada

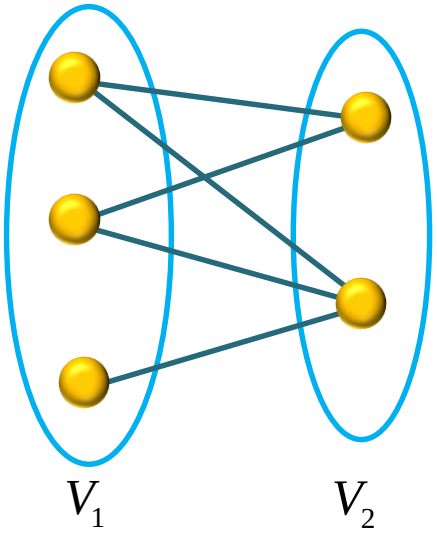
Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

University of Strathclyde

Network Bipartivity

bipartite graph



$$V = V_1 \cup V_2$$

$$A = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}$$

Adjacency matrix

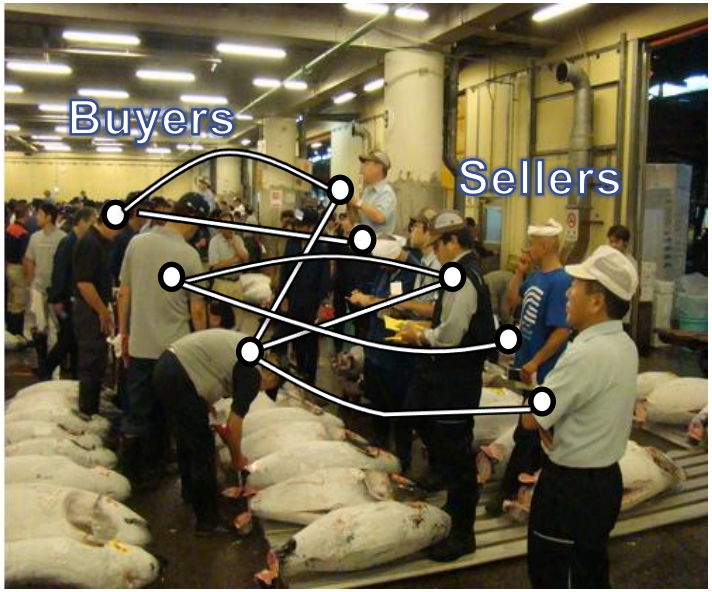
$$\lambda_j = -\lambda_{n-j+1}$$

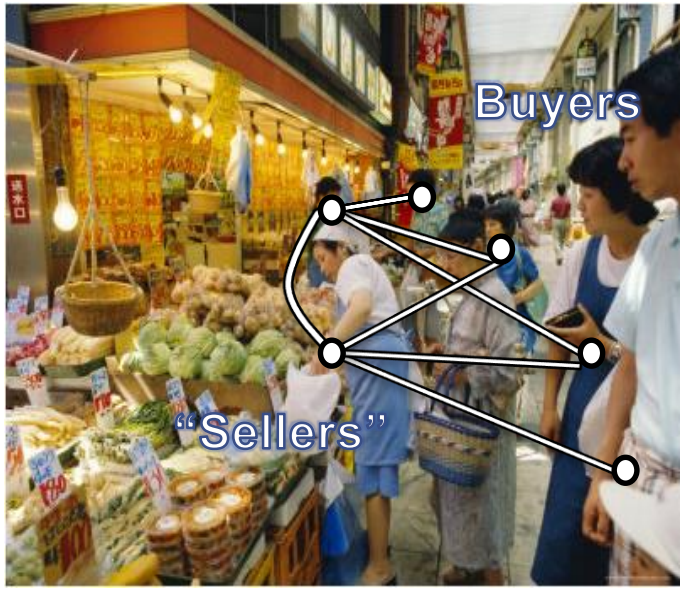
spectrum

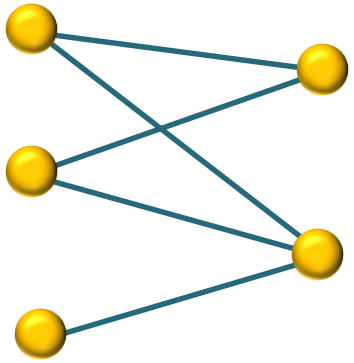
(see Chapters 2 & 5)

University of Strathclyde

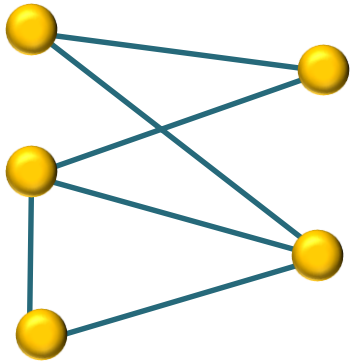
Network Bipartivity







Bipartite Network



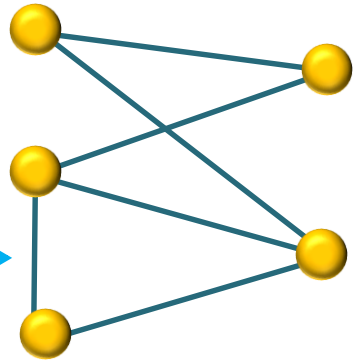
How much bipartite?

University of Strathclyde

Network Bipartivity

Computational approach

$$b = 1 - \frac{m_{fr}}{m}$$



Problem:

Need to find the best bipartition of the network to identify the number of frustrating edges. The problem is NP-complete.

University of Strathclyde

Network Bipartivity

Algebraic approach

$$Bipartivity \Leftrightarrow No\ odd\ \ cycles$$

$$Bipartivity \Leftrightarrow Tr\left(\sinh\left(A\right)\right) = 0$$

University of Strathclyde

Network Bipartivity

Algebraic approach

$$Tr\left(\exp\left(A\right)\right) = Tr\left(\sinh\left(A\right)\right) + Tr\left(\cosh\left(A\right)\right)$$

$$Bipartivity \Leftrightarrow Tr\left(\exp\left(A\right)\right) = Tr\left(\cosh\left(A\right)\right)$$

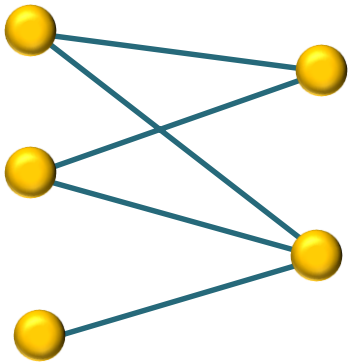
$$b_s = \frac{Tr\left(\cosh\left(A\right)\right)}{Tr\left(\exp\left(A\right)\right)} = \frac{\sum_{j=1}^{n}\cosh\left(\lambda_j\right)}{\sum_{j=1}^{n}\exp\left(\lambda_j\right)} = \frac{EE_{even}}{EE}$$

University of Strathclyde

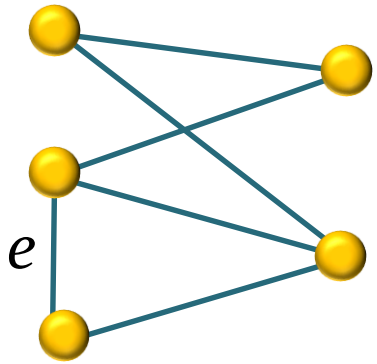
Network Bipartivity

Theorem:

$$b_s(G) = \frac{EE_{even}}{EE} \geq \frac{EE_{even} + a}{EE + a + b} = b_s(G + e)$$



G



$G + e$

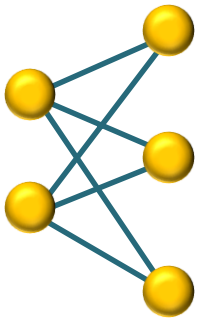
a **contribution to even CWs**

b **contribution to odd CWs**

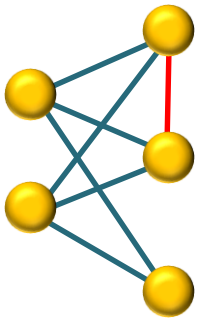
University of Strathclyde

Network Bipartivity

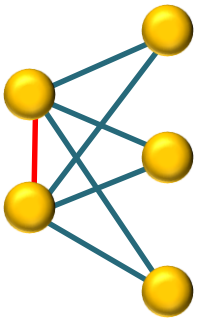
Algebraic approach



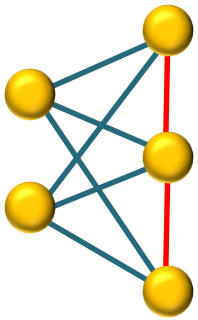
$b_s = 1.000$



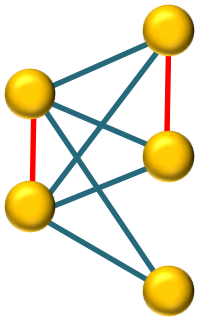
$b_s = 0.829$



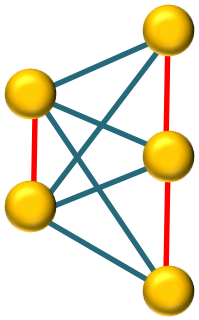
$b_s = 0.769$



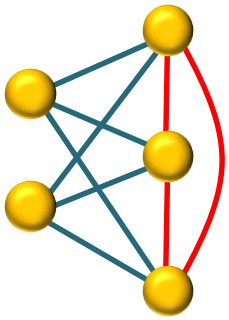
$b_s = 0.721$



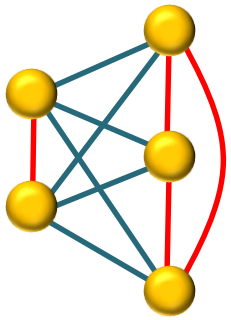
$b_s = 0.692$



$b_s = 0.645$



$b_s = 0.645$



$b_s = 0.597$

University of Strathclyde

Network Bipartivity

(Another) algebraic approach

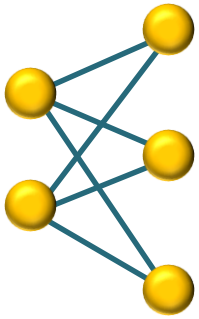
$$b_e = \frac{\sum_{j=1}^{n} \cosh\left(\lambda_j\right) - \sum_{j=1}^{n} \sinh\left(\lambda_j\right)}{\sum_{j=1}^{n} \cosh\left(\lambda_j\right) + \sum_{j=1}^{n} \sinh\left(\lambda_j\right)}$$

$$b_e = \frac{Tr\left(\exp\left(-A\right)\right)}{Tr\left(\exp\left(A\right)\right)} = \frac{\sum_{j=1}^{n} \exp\left(-\lambda_j\right)}{\sum_{j=1}^{n} \exp\left(\lambda_j\right)}$$

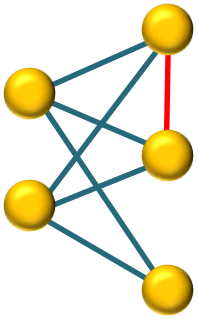
University of Strathclyde

Network Bipartivity

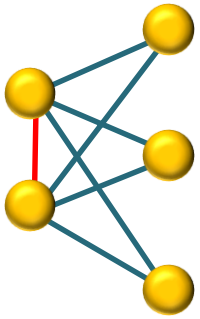
(Another) algebraic approach



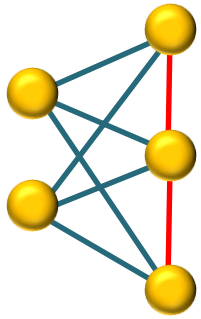
$b_s = 1.000$



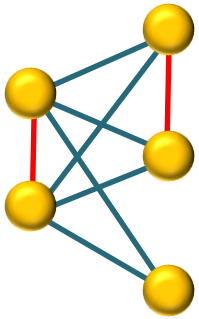
$b_s = 0.658$



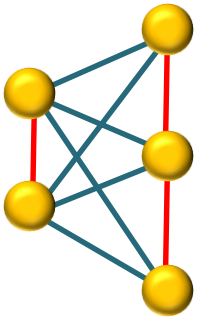
$b_s = 0.538$



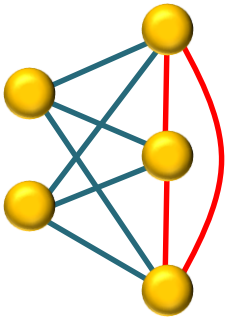
$b_s = 0.462$



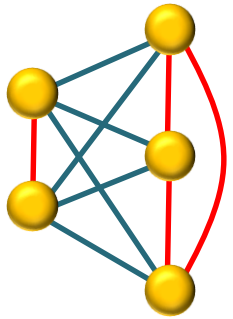
$b_s = 0.383$



$b_s = 0.289$



$b_s = 0.289$

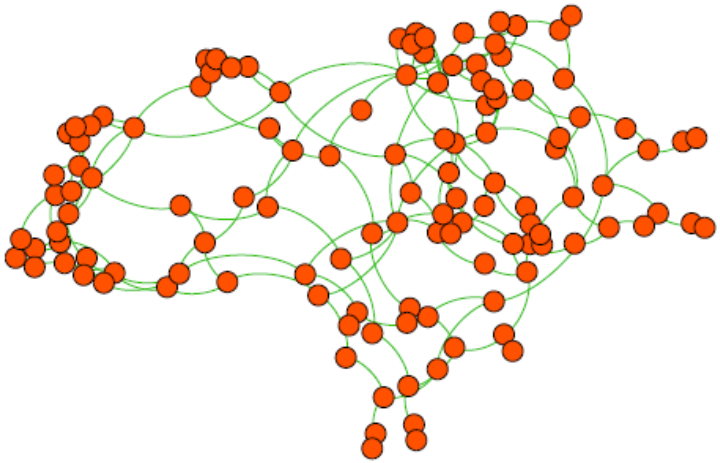


$b_s = 0.194$

Network Bipartivity

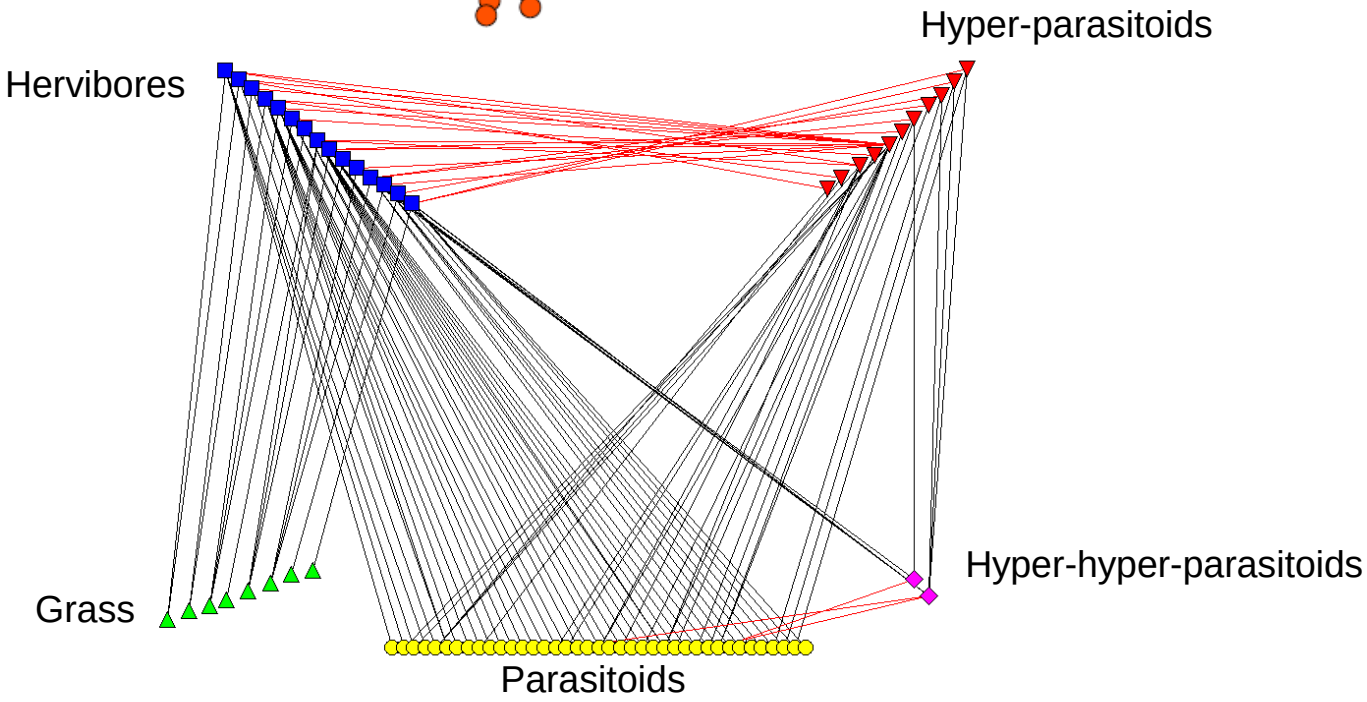
Example

A grassland food wen in England

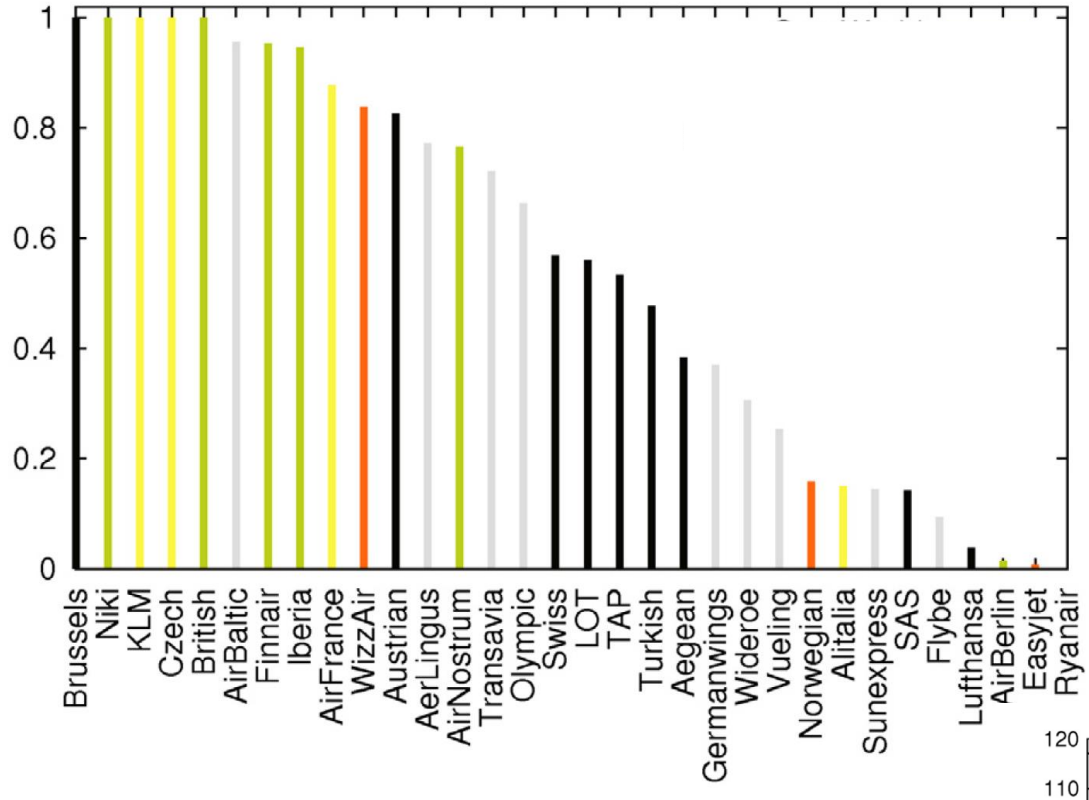


$$b_s = 0.766$$

$$b_e = 0.532$$



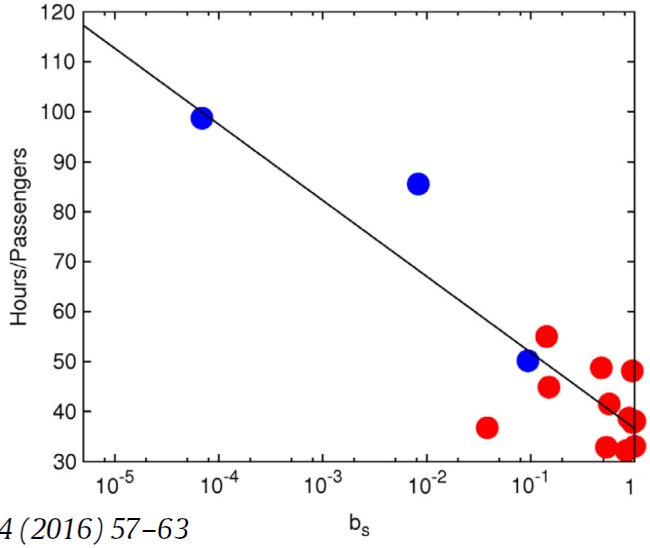
Bipartivity & Airlines Efficiency



One World
Sky Team
Star Alliance
Code-sharing LCC
Independent LCC

low cost

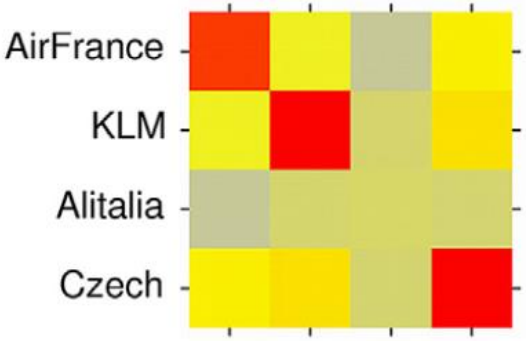
traditional



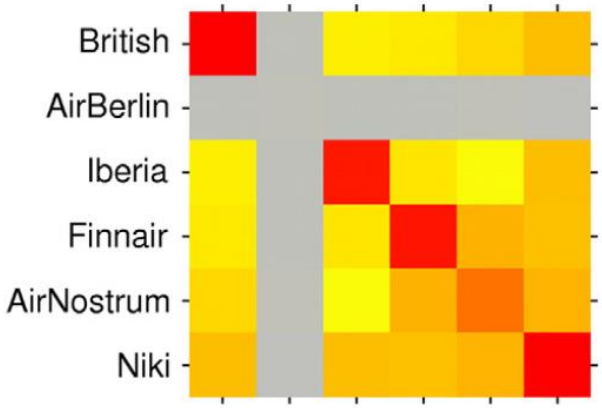
E. Estrada, J. Gómez-Gardeñes / Physica D 323–324 (2016) 57–63

Bipartivity & Airlines Efficiency

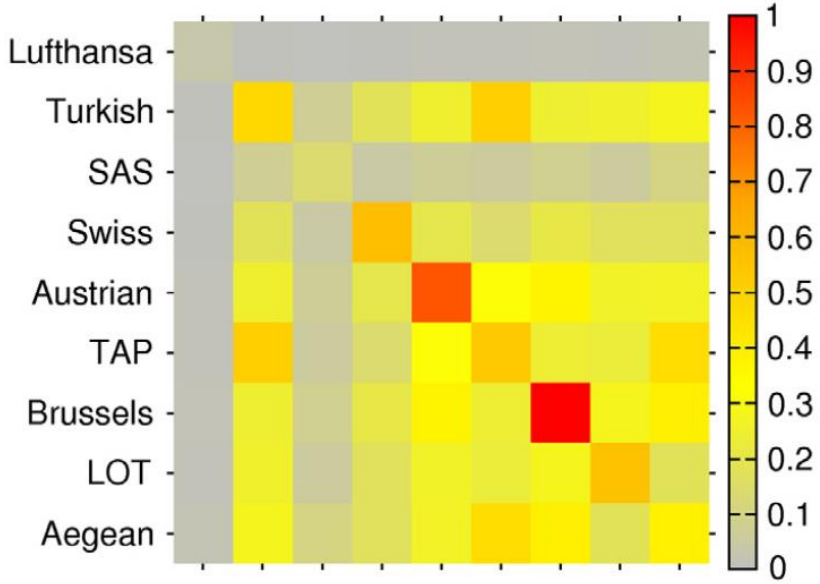
Sky team



One World



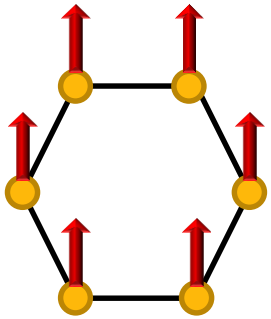
Star Alliance



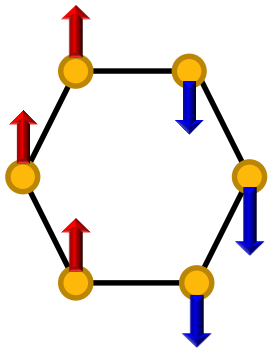
E. Estrada, J. Gómez-Gardeñes / Physica D 323–324 (2016) 57–63

University of Strathclyde

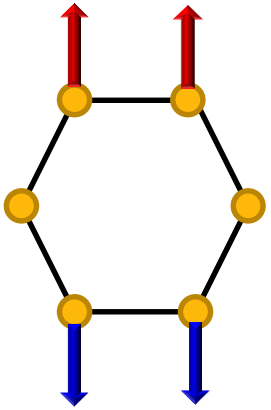
Finding Bipartitions



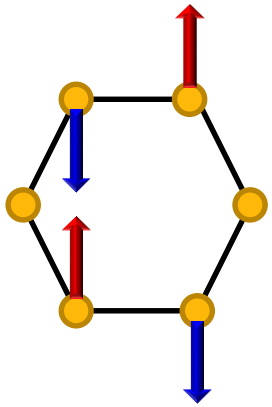
$j = 1$



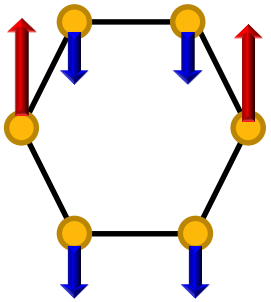
$j = 2$



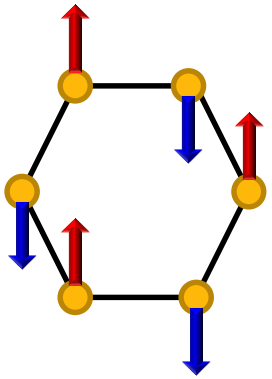
$j = 3$



$j = 4$



$j = 5$



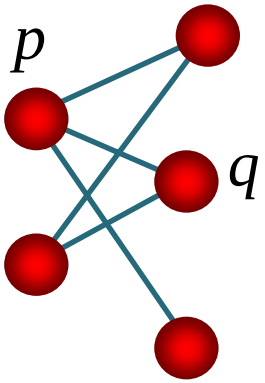
$j = 6$

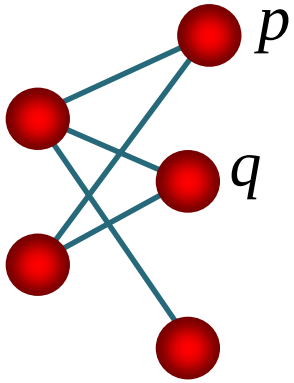
(see Chapter 21)

University of Strathclyde

Finding Bipartitions

$$\hat{G}_{pq} = \left[\exp\left(-A\right)\right]_{pq} = \left[\cosh\left(A\right) - \sinh\left(A\right)\right]_{pq}$$





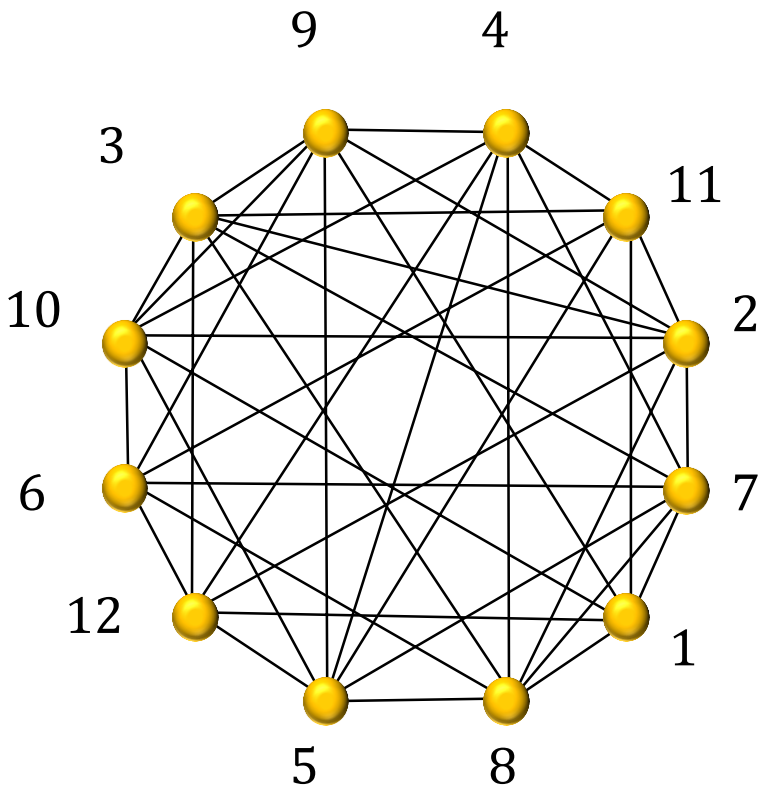
$$\hat{G}_{pq} = \left[-\sinh\left(A\right)\right]_{pq} < 0$$

$$\hat{G}_{pq} = \left[\cosh\left(A\right)\right]_{pq} > 0$$

$$\hat{G}_{pq} \begin{cases} < 0 & p \text{ and } q \text{ in different partitions} \\ > 0 & p \text{ and } q \text{ in the same partitions} \end{cases}$$

University of Strathclyde

Finding Bipartitions



Bipartite?

University of Strathclyde

Finding Bipartitions

1) Build the G(-1) matrix

$$\hat{G} = \begin{pmatrix}
10.1 & 6.83 & 7.09 & 7.11 & 6.66 & 8.19 & -5.62 & -7.98 & -7.29 & -7.28 & -9.02 & -8.73 \\
6.83 & 8.03 & 4.17 & 5.87 & 5.43 & 6.72 & -6.31 & -4.32 & -6.56 & -5.65 & -7.28 & -7.04 \\
7.09 & 4.17 & 8.05 & 5.62 & 5.50 & 6.73 & -5.42 & -6.62 & -4.30 & -6.56 & -7.29 & -7.05 \\
7.11 & 5.87 & 5.62 & 8.17 & 3.93 & 6.80 & -5.68 & -5.96 & -6.62 & -4.32 & -7.98 & -7.10 \\
6.66 & 5,43 & 5.50 & 3.93 & 7.37 & 6.36 & -5.39 & -5.68 & -5.42 & -6.31 & -5.62 & -6.77 \\
8.19 & 6.72 & 6.73 & 6.80 & 6.36 & 9.38 & -6.77 & -7.10 & -7.05 & -7.04 & -8.73 & -7.28 \\
-5.62 & -6.31 & -5.42 & -5.68 & -5.39 & -6.77 & 7.37 & 3.93 & 5.50 & 5.43 & 6.66 & 6.36 \\
-7.98 & -4.32 & -6.62 & -5.96 & -5.68 & -7.10 & 3.93 & 8.17 & 5.62 & 5.87 & 7.11 & 6.80 \\
-7.29 & -6.56 & -4.30 & -6.62 & -5.42 & -7.05 & 5.50 & 5.62 & 8.05 & 4.17 & 7.09 & 6.73 \\
-7.28 & -5.65 & -6.56 & -4.32 & -6.31 & -7.04 & 5.43 & 5.87 & 4.17 & 8.03 & 6.83 & 6.72 \\
-9.02 & -7.28 & -7.29 & -7.98 & -5.62 & -8.73 & 6.66 & 7.11 & 7.09 & 6.83 & 10.1 & 8.19 \\
-8.73 & -7.04 & -7.05 & -7.10 & -6.77 & -7.28 & 6.36 & 6.80 & 6.73 & 6.72 & 8.19 & 9.38
\end{pmatrix}$$

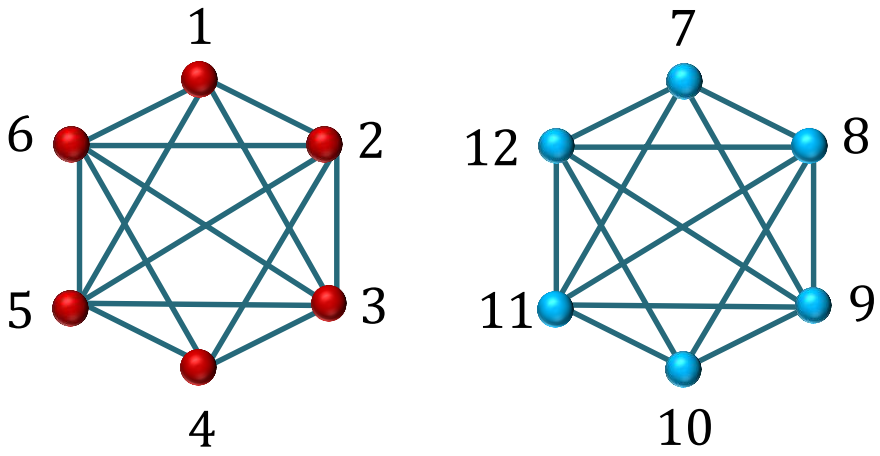
Finding Bipartitions

2) Dichotomise G(-1). Take the diagonal as zeros

$$\tilde{G} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

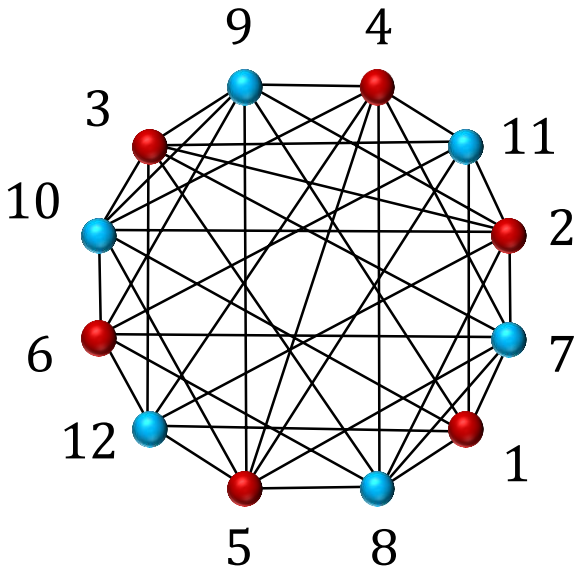
University of Strathclyde

3) Construct the anti-communicability graph



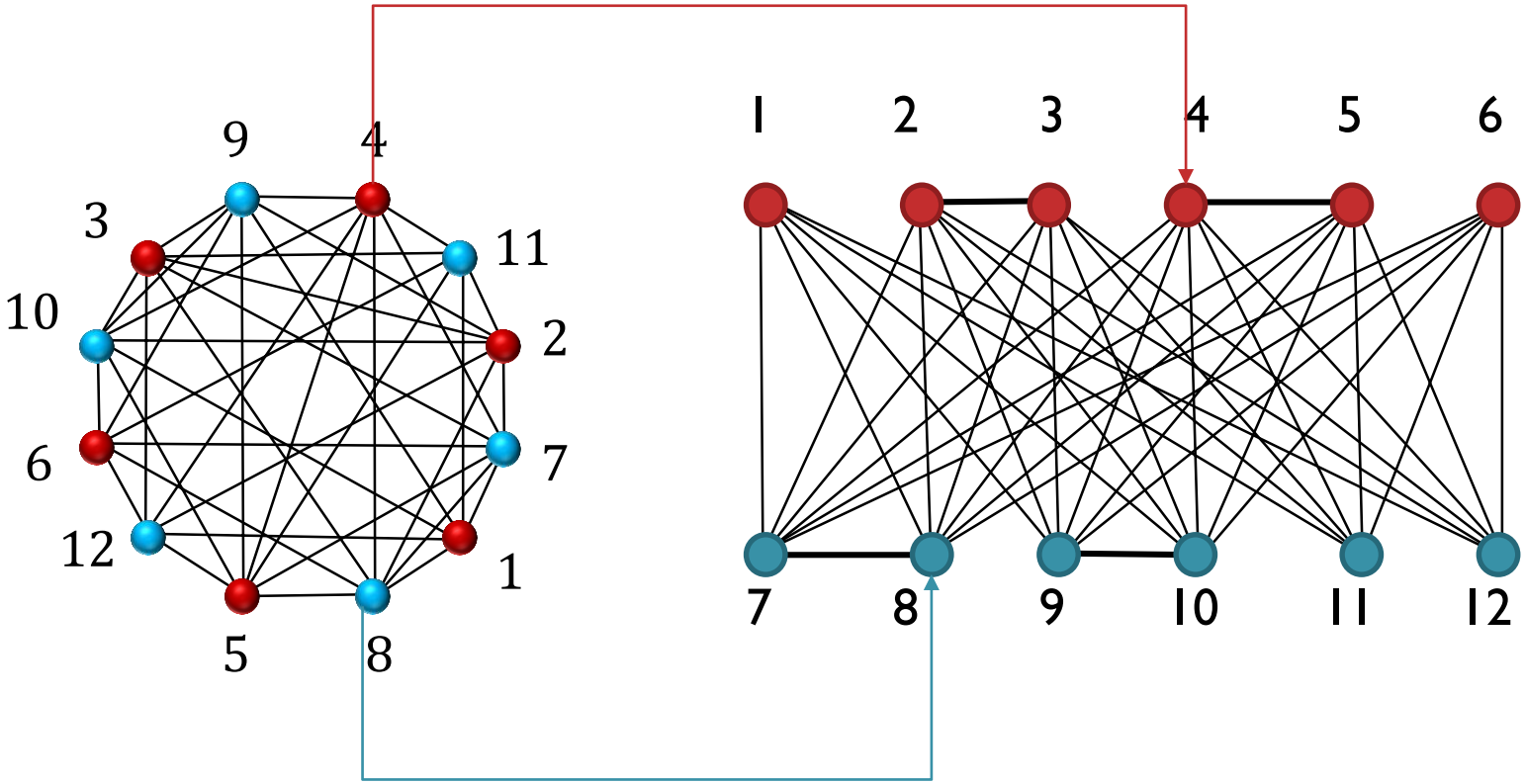
$$C_1 = \{1, 2, 3, 4, 5, 6\}$$

$$C_2 = \{6, 7, 8, 9, 10, 11, 12\}$$



University of Strathclyde

4) Represent the partitions of the network

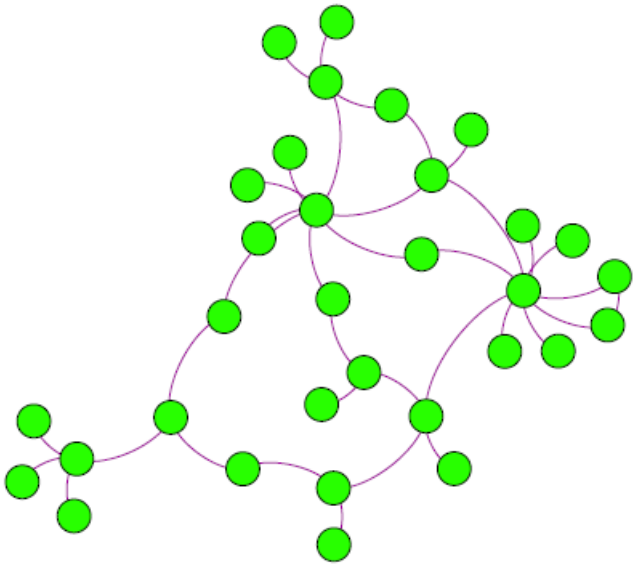


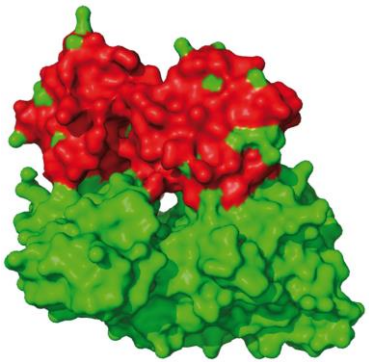
University of Strathclyde

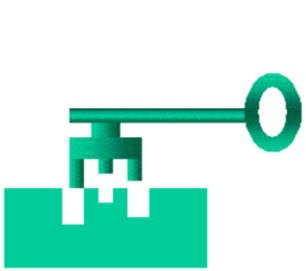
Finding Bipartitions

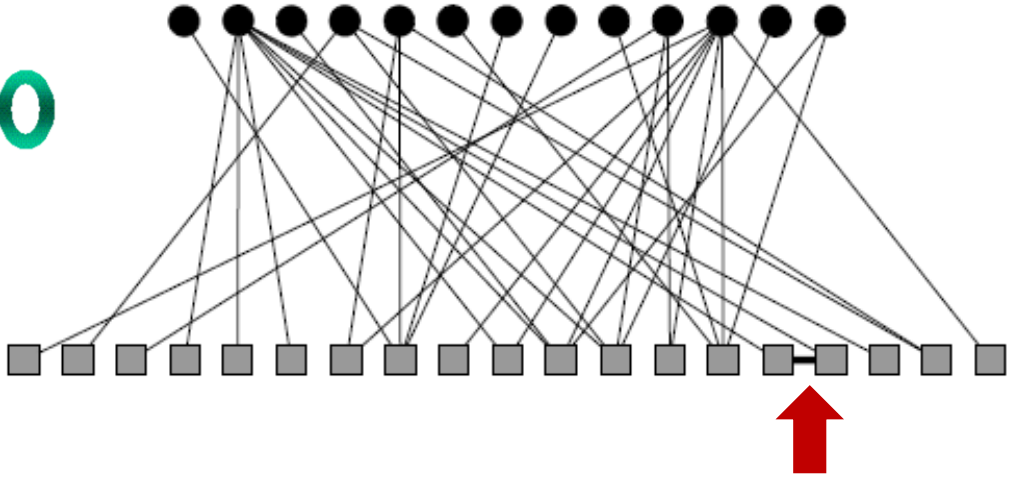
Example

Protein-protein interaction network of *A. fulgidus*







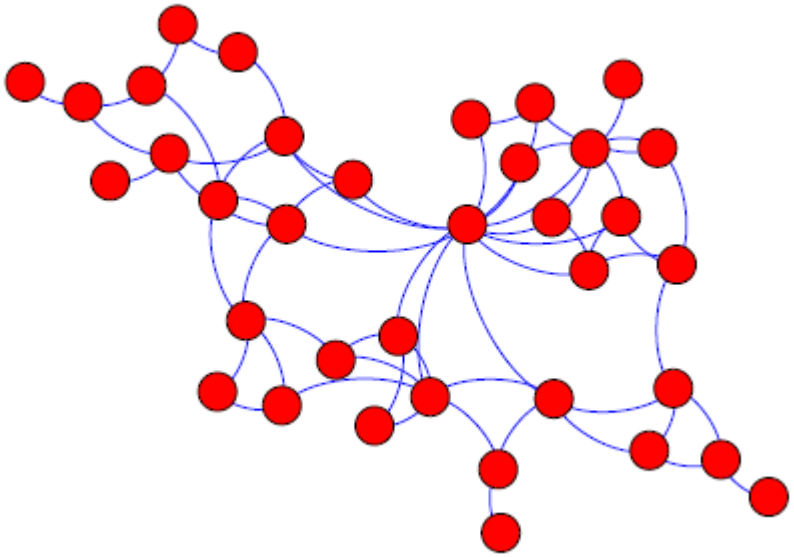


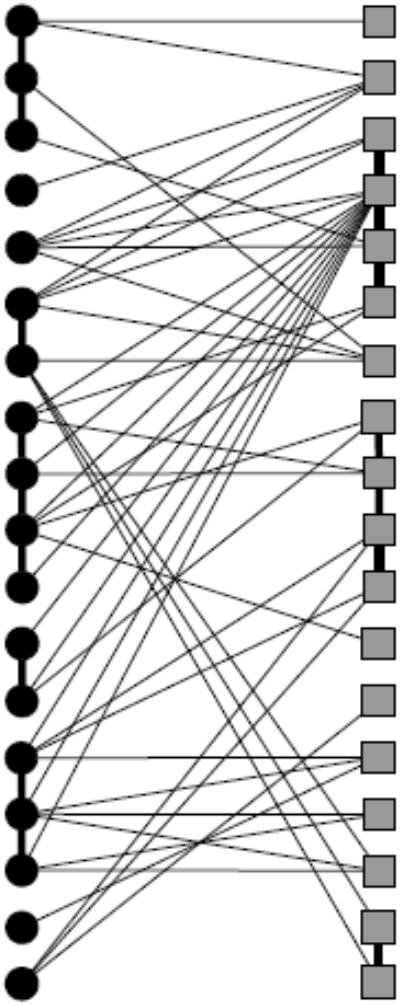
University of Strathclyde

Finding Bipartitions

Example

Social network in a sawmill





LECTURE 3: BASED ON CHAPTER 18 & 21

Degree-degree correlations. Statistical vs. Combinatorial Approaches

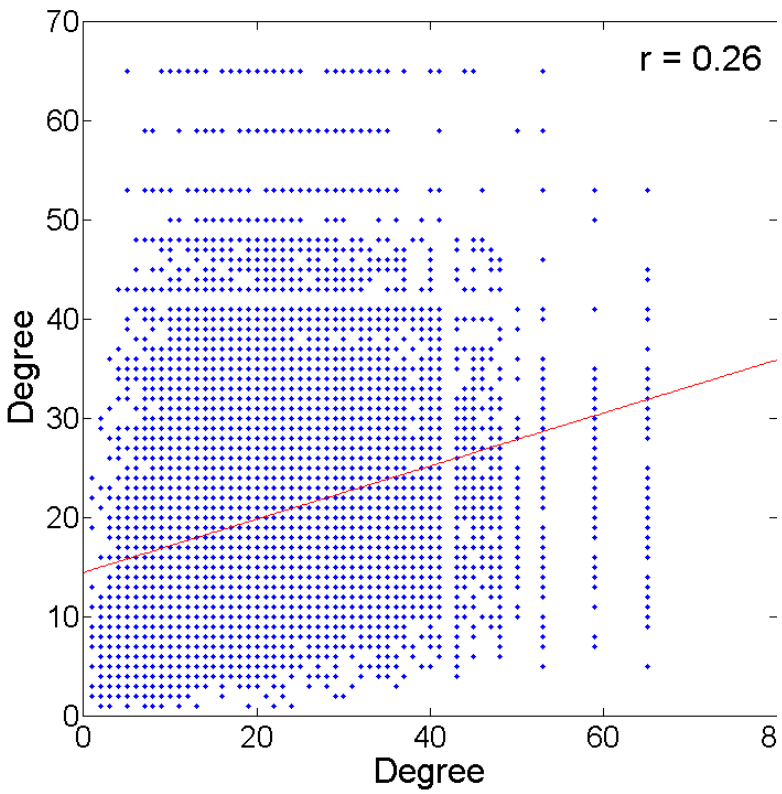
Ernesto Estrada

Department of Mathematics & Statistics
University of Strathclyde
Glasgow, UK

www.estradalab.org

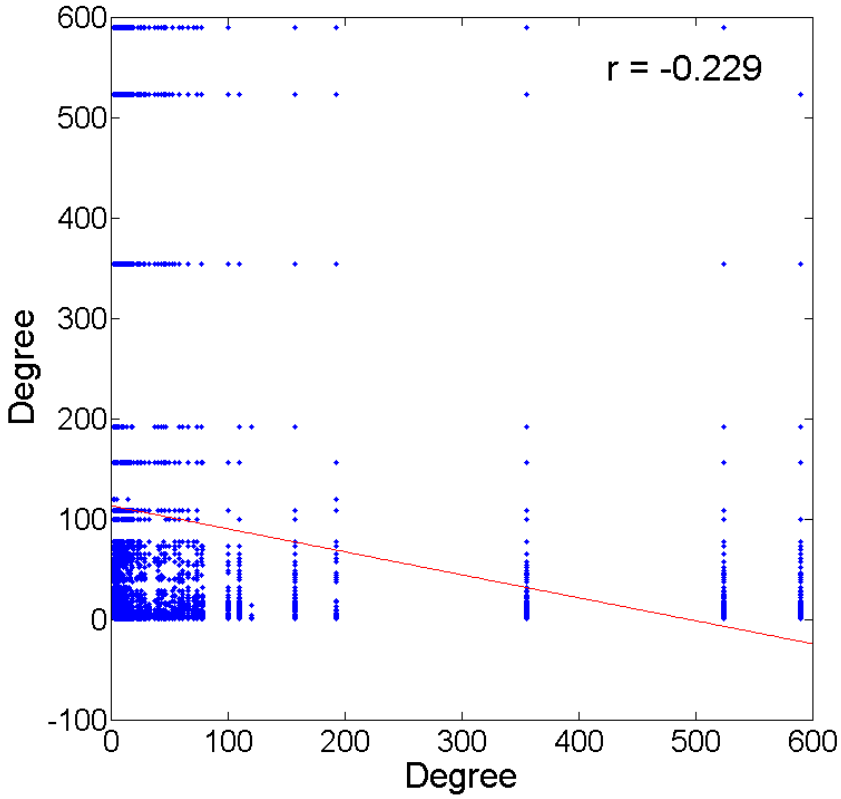
Degree-degree correlations

SOCIAL NETWORK



Assortative

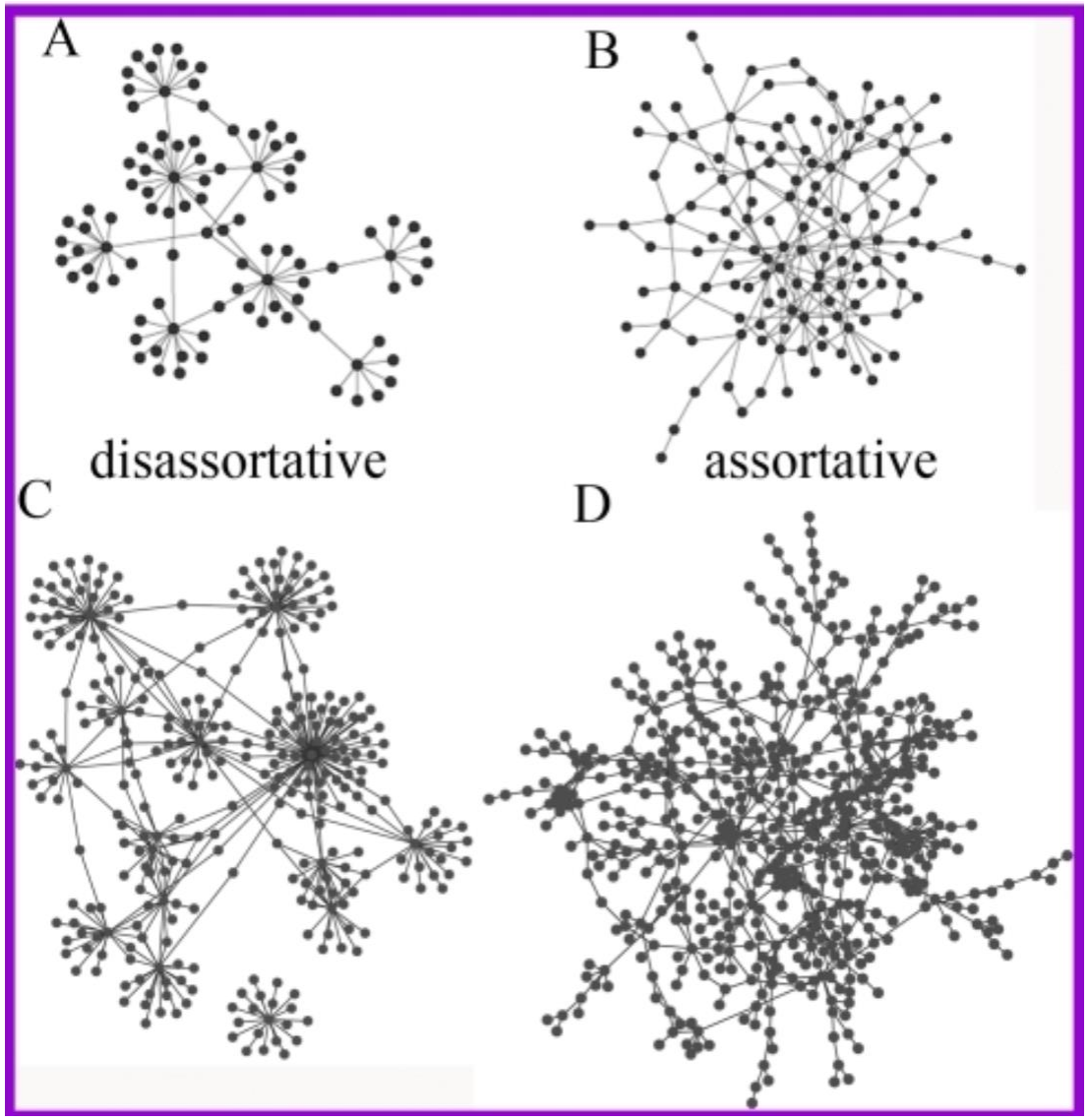
INTERNET



Disassortative

University of Strathclyde

Degree-degree correlations



University of Strathclyde

Degree-degree correlations

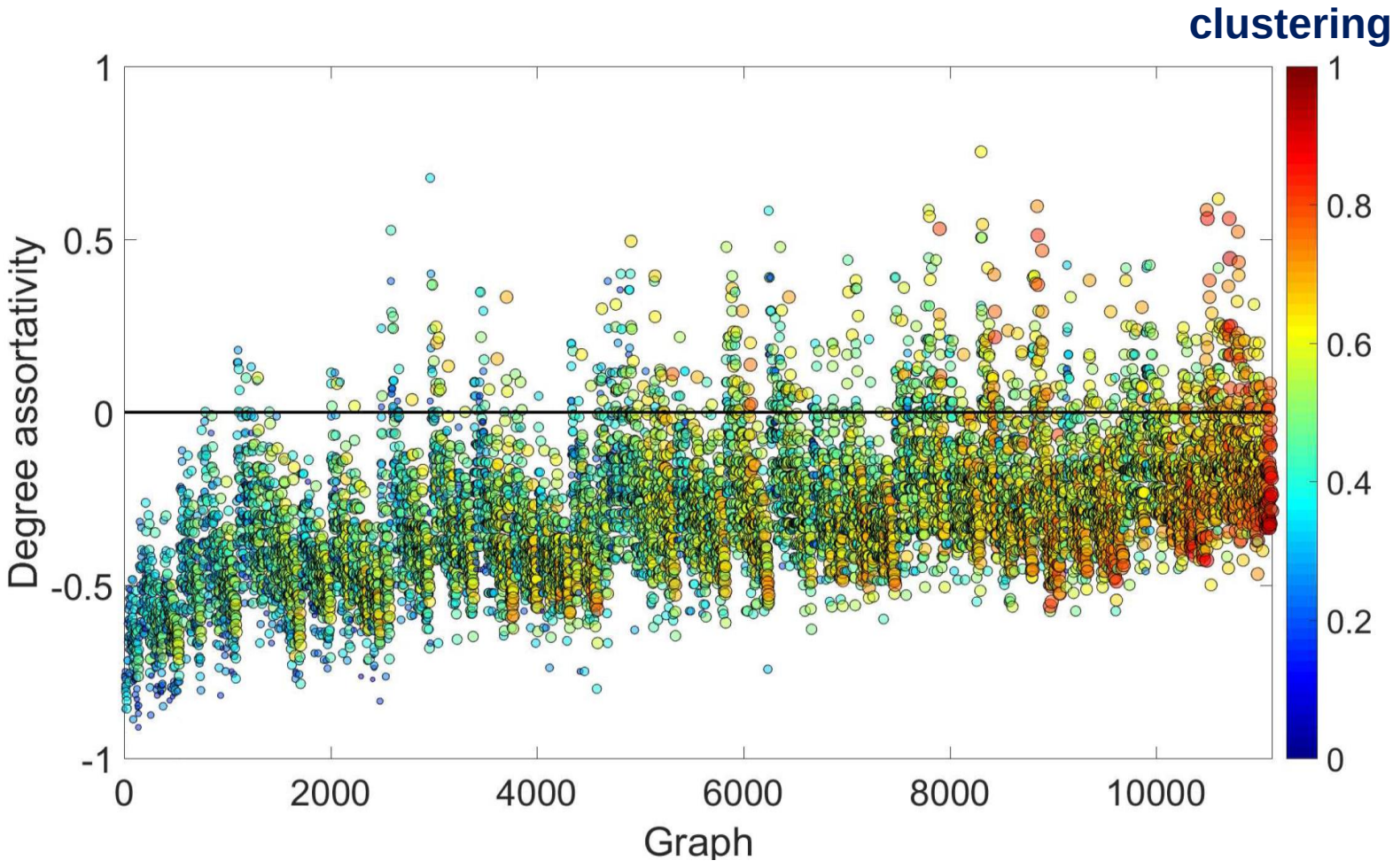
$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

$r > 0$ ***Assortative***

$r < 0$ ***Disassortative***

University of Strathclyde

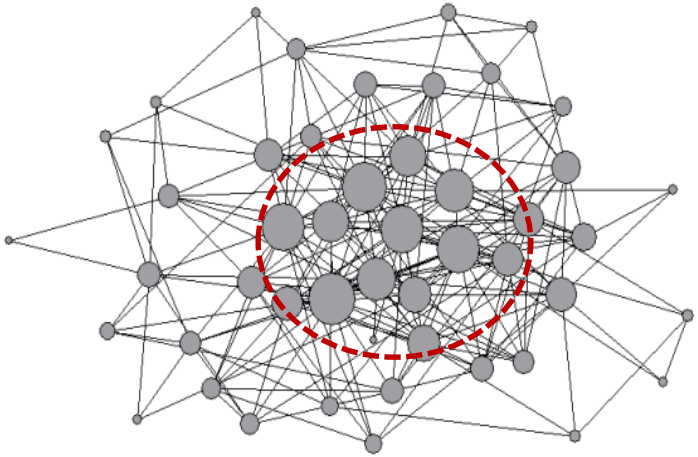
Degree-degree correlations



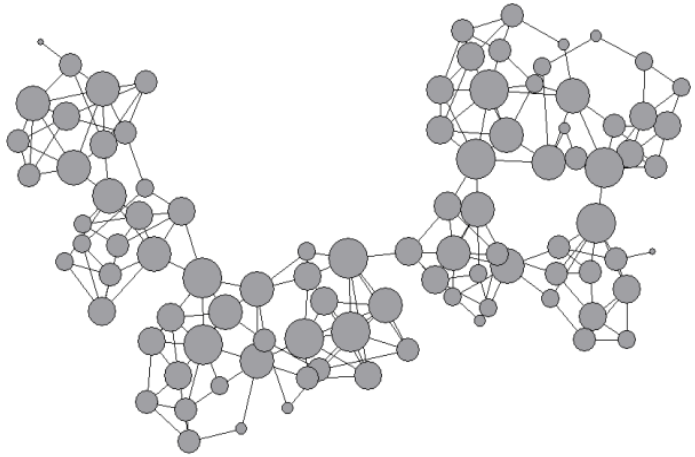
11,117 connected graphs with 8 nodes

University of Strathclyde

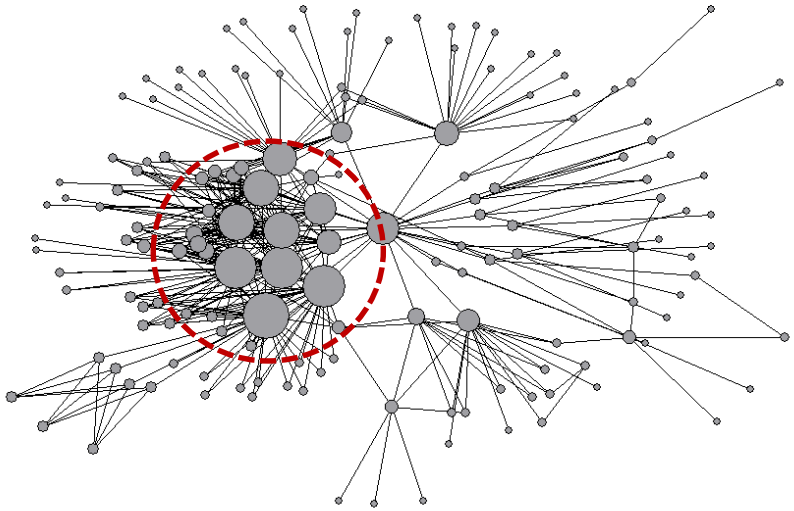
Degree-degree correlations



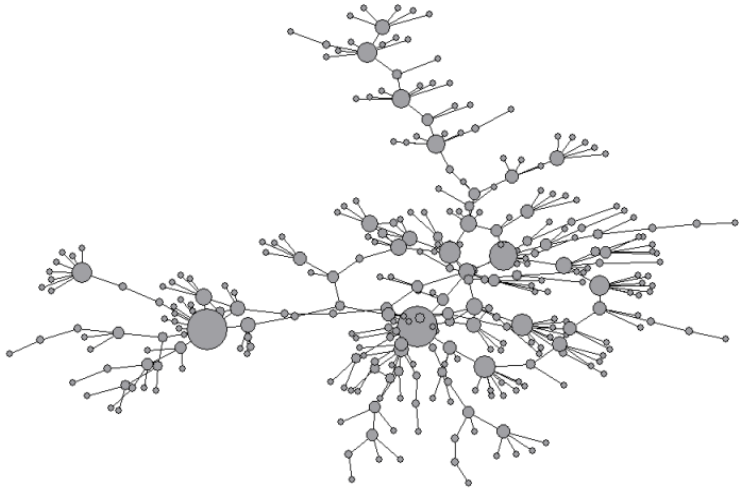
$r = 0.118$



$r = 0.129$



$r = -0.304$

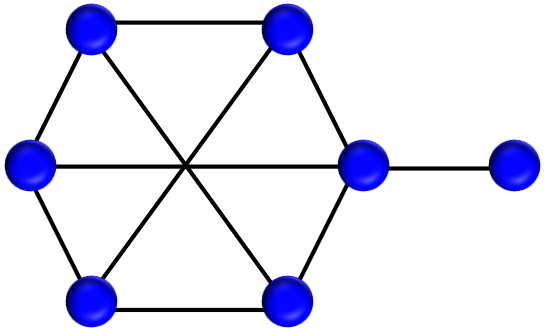


$r = -0.277$

University of Strathclyde

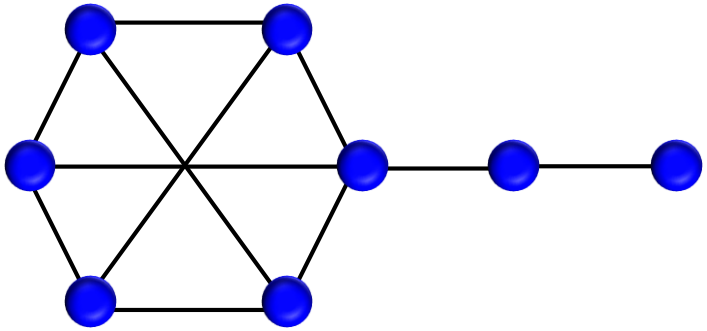
I. Degree-degree Correlations

Disassortative



r = -0.538

Assortative



r = 0.200

University of Strathclyde

Degree-degree correlations

$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

$$\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) = \sum_t k_t^3 \qquad \left(\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2 = \left(\sum_t k_t^2\right)^2$$

$$\frac{1}{2m}\left[\sum_t k_t^3 - \frac{1}{2m}\left(\sum_t k_t^2\right)^2\right]$$

$$\frac{1}{4m^2}\left[\left(\sum_t k_t\right)\left(\sum_t k_t^3\right) - \left(\sum_t k_t^2\right)^2\right]$$

Degree-degree correlations

We need to check:

$$\left(\sum_i k_i\right)\left(\sum_i k_i^3\right)-\left(\sum_i k_i^2\right)^2 \overset{?}{\geq} 0$$

Then, we write:

$$\sum_{i,j} k_i k_j \left(k_i^2 + k_j^2\right) - 2\sum_{i,j}\left(k_i k_j\right)^2 \overset{?}{\geq} 0$$

$$\sum_{i,j} k_i k_j \left(k_i^2 + k_j^2 - 2k_i k_j\right) \overset{?}{\geq} 0$$

$$\sum_{i,j} k_i k_j \left(k_i - k_j\right)^2 \geq 0.$$

Which proves that the denominator is never negative.

Degree-degree correlations

$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

Let us analyse every part of the Pearson coefficient:

$$\sum_{(i,j)\in E}\left(k_i + k_j\right) = \sum_t k_t^2$$

$$\sum_t k_t^2 = \sum_t \left(k_t^2 - k_t + k_t\right)$$

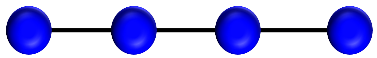
$$= \sum_t \left(k_t\left(k_t - 1\right)\right) + \sum_t k_t$$

$$= 2\left|P_2\right| + 2m;$$

$$\left(\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2 = \left(2\left|P_2\right| + 2m\right)^2$$

Degree-degree correlations

$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

Let us write $P_3 =$  as:

$$|P_3| = \sum_{(i,j)\in E}\left(k_i - 1\right)\left(k_j - 1\right) - 3|C_3|$$

$$= \sum_{(i,j)\in E}\left(k_i k_j\right) - \sum_{(i,j)\in E}\left(k_i + k_j\right) + m - 3|C_3|$$

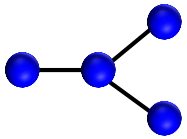
Thus:

$$\sum_{(i,j)\in E}\left(k_i k_j\right) = \sum_{(i,j)\in E}\left(k_i + k_j\right) - m + 3|C_3| + |P_3|$$

$$= m + 2|P_2| + |P_3| + 3|C_3|$$

Degree-degree correlations

$$r = \frac{\frac{1}{m}\sum_{(i,j)\in E} k_i k_j - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}{\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) - \left(\frac{1}{2m}\sum_{(i,j)\in E}\left(k_i + k_j\right)\right)^2}$$

Let us write $S_{1,3} =$  as:

$$\left|S_{1,3}\right| = \sum_t \binom{k_t}{3} = \frac{1}{6}\sum_i k_t\left(k_t - 1\right)\left(k_t - 2\right)$$

$$= \frac{1}{6}\sum_t k_t^3 - \frac{1}{2}\sum_t k_t^2 + \frac{1}{3}\sum_t k_t$$

Then:

$$\sum_t k_t^3 = 6\left|S_{1,3}\right| + 3\sum_t k_t^2 - 2\sum_t k_t = 6\left|S_{1,3}\right| + 6\left|P_2\right| + 2m$$

Consequently,

$$\sum_{(i,j)\in E}\left(k_i^2 + k_j^2\right) = 6\left|S_{1,3}\right| + 6\left|P_2\right| + 2m$$

Degree-degree correlations

We can the write,

$$r = \frac{\frac{1}{m}\left(m + 2|P_2| + |P_3| + 3|C_3|\right) - \frac{1}{4m^2}\left(2|P_2| + 2m\right)^2}{\frac{1}{2m}\left(6|S_{1,3}| + 6|P_2| + 2m\right) - \frac{1}{4m^2}\left(2|P_2| + 2m\right)^2}$$

Using $|P_{r/s}| = |P_r| / |P_s|$ we get

$$r = \frac{|P_2|\left(|P_{3/2}| + \frac{3|C_3|}{|P_2|} - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

And finally,

$$r = \frac{|P_2|\left(|P_{3/2}| + C - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

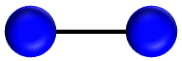
Degree-degree correlations

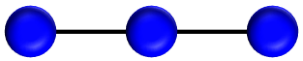
Resuming:

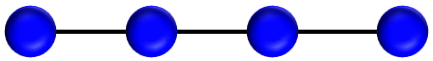
$$r = \frac{\left|P_2\right|\left(\left|P_{3/2}\right| + C - \left|P_{2/1}\right|\right)}{3\left|S_{1,3}\right| + \left|P_2\right|\left(1 - \left|P_{2/1}\right|\right)}$$

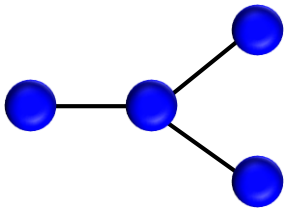
$$\left|P_{r/s}\right| = \left|P_r\right| / \left|P_s\right|$$

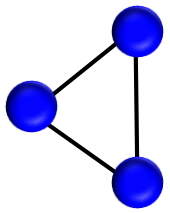
$$C = 3\left|C_3\right| / \left|P_2\right|$$

$P_1 =$ 

$P_2 =$ 

$P_3 =$ 

$S_{1,3} =$ 

$C_3 =$ 

Estrada, E., Combinatorial study of degree assortativity in networks. *Physical Review E* 84 **2011**, 047101

University of Strathclyde

Degree-degree correlations

Assortative

$$r > 0$$

$$\left|P_{2/1}\right| < \left|P_{3/2}\right| + C$$

Disassortative

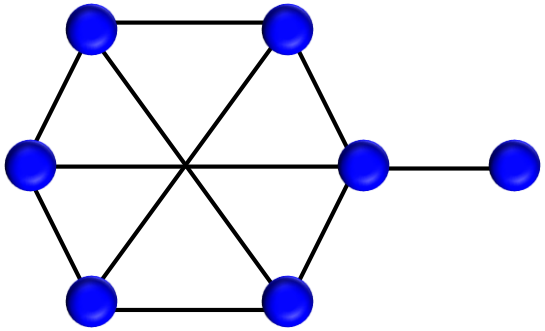
$$r < 0$$

$$\left|P_{2/1}\right| > \left|P_{3/2}\right| + C$$

Degree-degree correlations

DISASSORTATIVE

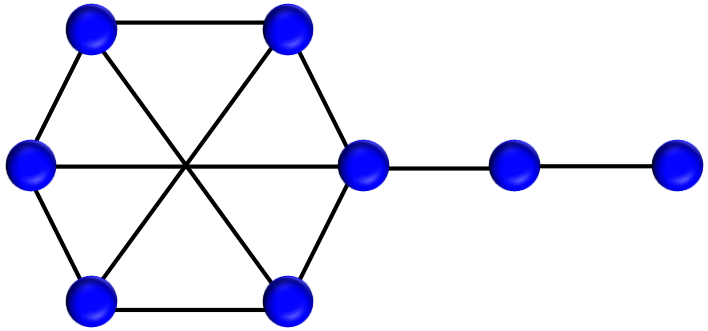
$$\left|P_{2/1}\right| > \left|P_{3/2}\right| + C$$



r = -0.538

ASSORTATIVE

$$\left|P_{2/1}\right| < \left|P_{3/2}\right| + C$$



r = 0.200

Degree-degree correlations

$$r = \frac{|P_2|\left(|P_{3/2}| + C - |P_{2/1}|\right)}{3|S_{1,3}| + |P_2|\left(1 - |P_{2/1}|\right)}$$

TABLE I. Relative branching ($|P_{2/1}|$), transitivity (C), intermodular connectivity ($|P_{3/2}|$), and assortativity coefficient for real-world networks.

Network	$\lvert P_{2/1}\rvert$		$\lvert P_{3/2}\rvert$		C	r
Prison	4.25	<	(4.09	+	0.288)	
Protein residue	4.41	<	(4.45	+	0.417)	
St. Marks	10.54	<	(10.46	+	0.291)	
Geom	17.42	<	(22.09	+	0.224)	
Corporate	19.42	<	(20.60	+	0.498)	
Roget	9.55	<	(10.08	+	0.134)	
Jazz	127.30	<	(144.84	+	0.771)	
Zachary	6.77	>	(4.49	+	0.256)	
Drugs	14.58	>	(12.84	+	0.368)	
Transcription	12.51	>	(3.01	+	0.016)	
Bridge Brook	22.42	>	(17.31	+	0.191)	
USAir97	43.36	>	(36.97	+	0.396)	
Internet	91.00	>	(11.53	+	0.015)	


LOONEY TUNES
That's all