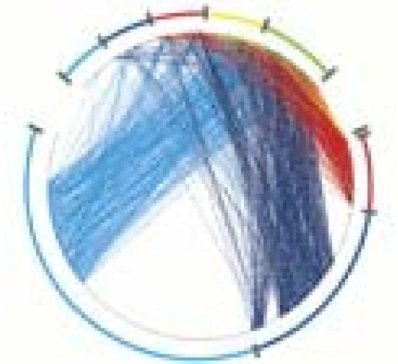
Lecture 5: Spectral properties & Network Motifs



Spectral properties of graphs

- A graph should be summarized in a number of meaningful metrics
- These metrics can be regarded as the features
- Visualization is a tool
- Spectral graph theory gives useful information on a graph
- It is indeed the eigenvalue decomposition of the adjacency or Laplacian matrices
- It links the graph to connectivity
- Connectivity of a graph is one of the most important issues
- Spectral analysis has applications is
  - Synchronization
  - Graph partitioning
  - Clustering
  - Coarse graining
  - …

Eigenvalue and eigenvectors

- Given a matrix **A**, **x** is the eigenvector and λ is the corresponding eigenvalue if $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$
  - **A** must be square and the determinant of $\mathbf{A} - \lambda\mathbf{I}$ must be equal to zero

$$\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = 0 \; ! \; (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = 0$$

    - Trivial solution is if $\mathbf{x} = 0$
    - The non trivial solution occurs when $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$
    - The eigenvalues and the corresponding eigenvectors are obtained by solving the above equation
- Are eigenvectors unique?
  - If **x** is an eigenvector, then $\beta\mathbf{x}$ is also an eigenvector

$$\mathbf{A}(\beta\mathbf{x}) = \beta(\mathbf{A}\mathbf{x}) = \beta(\lambda\mathbf{x}) = \lambda(\beta\mathbf{x})$$

Eigenvalue and eigenvectors

- An example for a 2 × 2 matrix:
- Expand the det(**A** - λ**I**) = 0 for a 2 × 2 matrix

$$\det\left(A-\lambda I\right)=\det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}-\lambda\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)=0$$

$$\det\begin{bmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{bmatrix}=0 \Rightarrow \left(a_{11}-\lambda\right)\left(a_{22}-\lambda\right)-a_{12}a_{21}=0$$

$$\lambda^2-\lambda\left(a_{11}+a_{22}\right)+\left(a_{11}a_{22}-a_{12}a_{21}\right)=0$$

- For a 2 × 2 matrix, this is a simple quadratic equation with two solutions (maybe complex)

$$\lambda=\left(a_{11}+a_{22}\right)\pm\sqrt{\frac{\left(a_{11}+a_{22}\right)^2}{4\left(a_{11}a_{22}-a_{12}a_{21}\right)}}$$

- This "characteristic equation" can be used to solve for **x**

Eigenvalue and eigenvectors

- Consider,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \Rightarrow \begin{cases} \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \\ \lambda^2 - (1+4)\lambda + (1 \cdot 4 - 2 \cdot 2) = 0 \\ \lambda^2 = (1+4)\lambda \Rightarrow \lambda = 0, \lambda = 5 \end{cases}$$

- The corresponding eigenvectors can be computed as

$$\lambda = 0 \Rightarrow \left[\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1x + 2y \\ 2x + 4y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

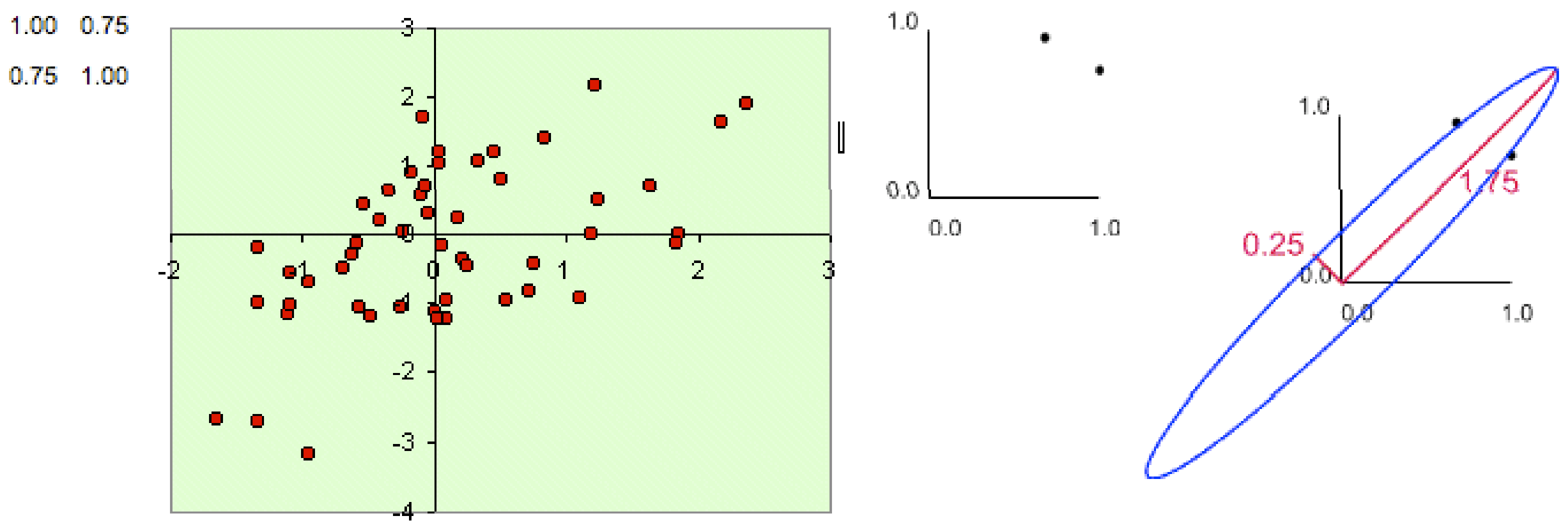
$$\lambda = 5 \Rightarrow \left[\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \right] \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -4x + 2y \\ 2x - 1y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

  - For $\lambda = 0$, one possible solution is $\mathbf{x} = (2, -1)$
  - For $\lambda = 5$, one possible solution is $\mathbf{x} = (1, 2)$

Physical interpretation

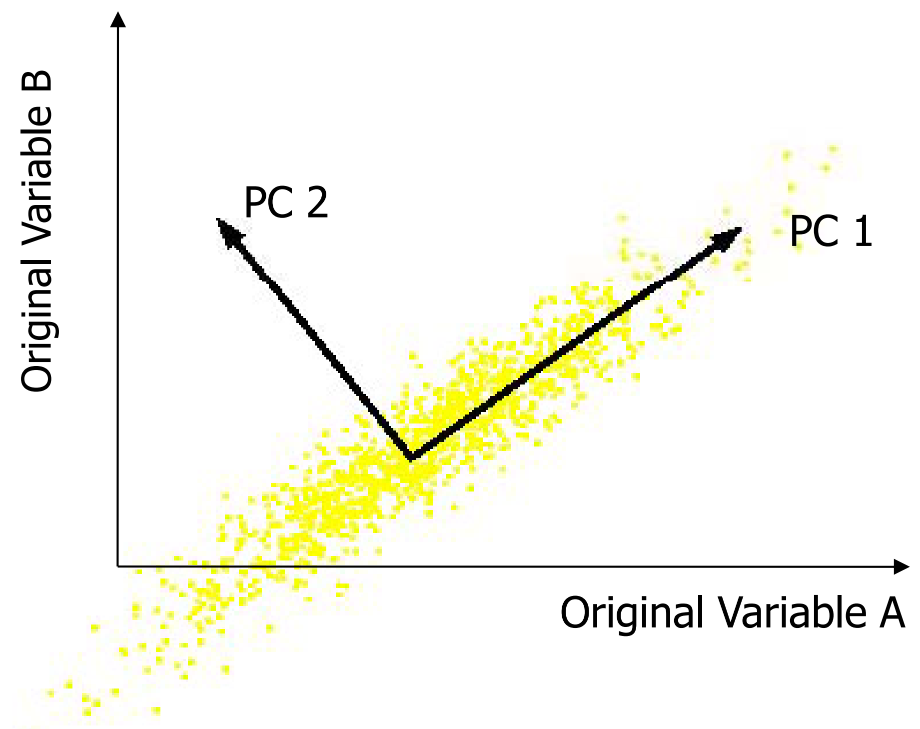
- Consider a covariance matrix, **A**, i.e., A = 1/n S S^T for some S

$$\mathbf{A} = \begin{bmatrix} 1 & .75 \\ .75 & 1 \end{bmatrix} \Rightarrow \lambda_1 = 1.75, \lambda_2 = 0.25$$



- Error ellipse with the major axis as the larger eigenvalue and the minor axis as the smaller eigenvalue

Physical interpretation



- Orthogonal directions of greatest variance in data
- Projections along PC1 (Principal Component) discriminate the data most along any one axis

Physical interpretation

- First principal component, which is in the direction of the eigenvector corresponding to the first eigenvalue, is the direction of greatest variability (covariance) in the data
- Second is the next orthogonal (uncorrelated) direction of greatest variability
  - So first remove all the variability along the first component, and then find the next direction of greatest variability
- And so on …
- Thus, each eigenvector provides the directions of data, represented by matrix **A**, variances in decreasing order of eigenvalues

Eigen/diagonal decomposition

- Let $\mathbf{S} \in \mathbb{R}^{m \times m}$ be a **square** matrix with ***m* linearly independent eigenvectors** (a "non-defective" matrix)
- **Theorem**: Exists an **eigen decomposition**

$$\mathbf{S} = \mathbf{U} \underbrace{\mathbf{\Lambda}}_{\textit{diagonal}} \mathbf{U}^{-1}$$

  - (matrix diagonalization theorem)

Unique for distinct eigen-values

- Columns of ***U*** are **eigenvectors** of ***S***
- Diagonal elements of $\mathbf{\Lambda}$ are **eigenvalues** of $\mathbf{S}$

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_m), \quad \lambda_i \geq \lambda_{i+1}$$

Eigen decomposition

Let ***U*** have the eigenvectors as columns: $U = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}$

Then, ***SU*** can be written

$$SU = S\begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} = \begin{bmatrix} \lambda_1 v_1 & \dots & \lambda_n v_n \end{bmatrix} = \begin{bmatrix} v_1 & \dots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \dots & \\ & & \lambda_n \end{bmatrix}$$

Thus ***SU=UΛ***, or ***U⁻¹SU=Λ***

And ***S=UΛU⁻¹***

Recall $S = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}; \lambda_1 = 1, \lambda_2 = 3.$ The eigenvectors $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ form $U = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

Inverting, we have $U^{-1} = \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

Recall $UU^{-1} = 1$.

Then, ***S=UΛU⁻¹*** = $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix}$

Eigen decomposition

Let's divide $\boldsymbol{U}$ (and multiply $\boldsymbol{U}^{-1}$) by $\sqrt{2}$

Then, $\boldsymbol{S}= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

$\boldsymbol{Q}$ $\quad\quad \Lambda \quad\quad$ $(\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{T})$

- If $\mathbf{S} \in \mathbb{R}^{m \times m}$ is a **symmetric** matrix:
- **Theorem**: Exists a (unique) **eigen decomposition**
- where $\boldsymbol{Q}$ is **orthogonal:**
  - $\boldsymbol{Q}^{-1} = \boldsymbol{Q}^{T}$
  - Columns of $\boldsymbol{Q}$ are normalized eigenvectors
  - Columns are orthogonal.
  - (everything is real)

Spectral decomposition theorem

- If **A** is a symmetric and positive definite k × k matrix ($\mathbf{x}^T\mathbf{A}\mathbf{x} > 0$) with λ_i ($\lambda_i > 0$) and $\mathbf{e}_i$, i = 1 ⋯ k being the k eigenvector and eigenvalue pairs, then

$$\underset{(k\times k)}{\mathbf{A}} = \lambda_1 \underset{(k\times 1)}{\mathbf{e}_1} \underset{(1\times k)}{\mathbf{e}_1^T} + \lambda_2 \underset{(k\times 1)}{\mathbf{e}_2} \underset{(1\times k)}{\mathbf{e}_2^T} \cdots + \lambda_k \underset{(k\times 1)}{\mathbf{e}_k} \underset{(1\times k)}{\mathbf{e}_k^T} \Rightarrow \underset{(k\times k)}{\mathbf{A}} = \sum_{i=1}^{k} \lambda_i \underset{(k\times 1)}{\mathbf{e}_i} \underset{(1\times k)}{\mathbf{e}_i^T} = \mathbf{P}\Lambda\mathbf{P}^T$$

$$\underset{(k\times k)}{\mathbf{P}} = [\mathbf{e}_1, \mathbf{e}_2 \cdots \mathbf{e}_k], \underset{(k\times k)}{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

  - This is also called the eigen decomposition theorem
- Any symmetric matrix can be reconstructed using its eigenvalues and eigenvectors

Spectral decomposition theorem

- Let **A** be a symmetric, positive definite matrix

$$A=\begin{bmatrix}2.2 & 0.4\\ 0.4 & 2.8\end{bmatrix} \Rightarrow \det(A-\lambda I)=0$$

$$\Rightarrow \lambda^2-5\lambda+(6.16-0.16)=(\lambda-3)(\lambda-2)=0$$

- The eigenvectors for the corresponding eigenvalues are
- Consequently,

$$\mathbf{e}_1^T=\left[\tfrac{1}{\sqrt{5}},\tfrac{2}{\sqrt{5}}\right],\mathbf{e}_2^T=\left[\tfrac{2}{\sqrt{5}},\tfrac{-1}{\sqrt{5}}\right]$$

$$A=\begin{bmatrix}2.2 & 0.4\\ 0.4 & 2.8\end{bmatrix}=3\begin{bmatrix}\tfrac{1}{\sqrt{5}}\\ \tfrac{2}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}\tfrac{1}{\sqrt{5}} & \tfrac{2}{\sqrt{5}}\end{bmatrix}+2\begin{bmatrix}\tfrac{2}{\sqrt{5}}\\ \tfrac{-1}{\sqrt{5}}\end{bmatrix}\begin{bmatrix}\tfrac{2}{\sqrt{5}} & \tfrac{-1}{\sqrt{5}}\end{bmatrix}$$

$$=\begin{bmatrix}0.6 & 1.2\\ 1.2 & 2.4\end{bmatrix}+\begin{bmatrix}1.6 & -0.8\\ -0.8 & 0.4\end{bmatrix}$$

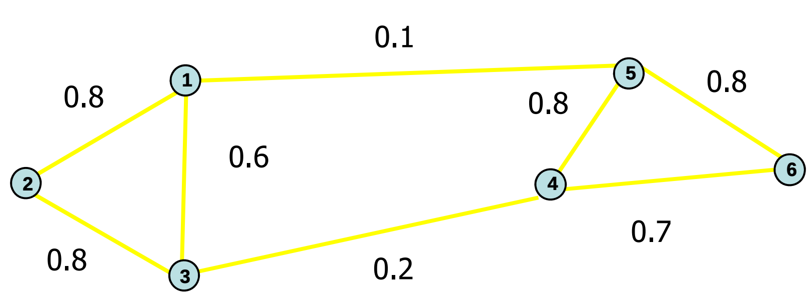
Spectral graph theory

- The eigenvalues and eigenvectors of a matrix provide global information about its structure
- Spectral Graph Theory
  - Analyse the "spectrum" of matrix representing a graph.
  - Spectrum: The eigenvectors of a graph, ordered by the magnitude(strength) of their corresponding eigenvalues.

$$\Lambda = \{\lambda_1, \lambda_2, ..., \lambda_n\}$$

Spectral graph theory

- Adjacency matrix (A)
  - *n x n* matrix
  - A = $[w_{ij}]$: edge weight between vertex x_i and x_j



	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	0.8	0.6	0	0.1	0
x_2	0.8	0	0.8	0	0	0
x_3	0.6	0.8	0	0.2	0	0
x_4	0	0	0.2	0	0.8	0.7
x_5	0.1	0	0	0.8	0	0.8
x_6	0	0	0	0.7	0.8	0

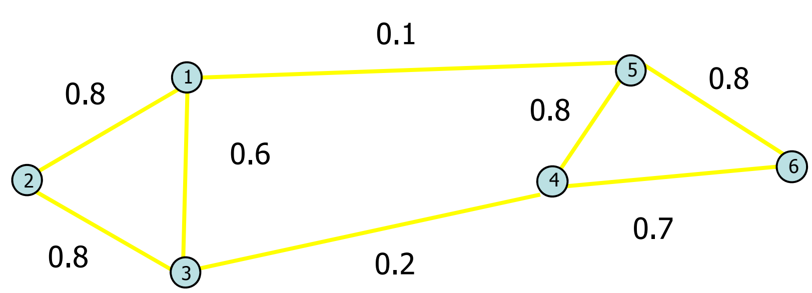
- Important properties:
  - Symmetric matrix

$\Rightarrow$ Eigenvalues are <u>real</u>

$\Rightarrow$ Eigenvector could span <u>orthogonal base</u>

Spectral graph theory

- Degree (or strength) matrix (D)
  - n x n diagonal matrix
  - $D(i,i) = \sum_j w_{ij}$: total weight of edges incident to vertex x_i

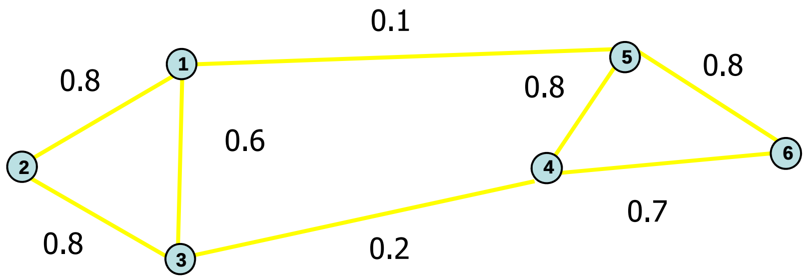


	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.5	0	0	0	0	0
x_2	0	1.6	0	0	0	0
x_3	0	0	1.6	0	0	0
x_4	0	0	0	1.7	0	0
x_5	0	0	0	0	1.7	0
x_6	0	0	0	0	0	1.5

- Important application:
  - Normalise adjacency matrix

Spectral graph theory

- Laplacian matrix (L): L = D – A
  - *n x n* symmetric matrix

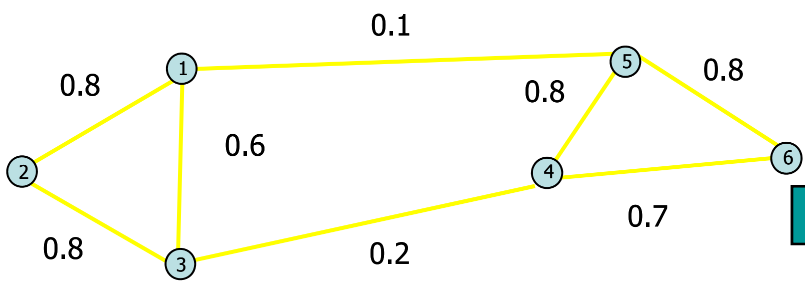


	x_1	x_2	x_3	x_4	x_5	x_6
x_1	1.5	-0.8	-0.6	0	-0.1	0
x_2	-0.8	1.6	-0.8	0	0	0
x_3	-0.6	-0.8	1.6	-0.2	0	0
x_4	0	0	-0.2	1.7	-0.8	-0.7
x_5	-0.1	0	0	-0.8	1.7	-0.8
x_6	0	0	0	-0.7	-0.8	1.5

- Important properties:
  - Eigenvalues are non-negative real numbers
  - Eigenvectors are real and orthogonal
  - Eigenvalues and eigenvectors provide an insight into the connectivity of the graph…

Spectral graph theory

- Normalized Laplacian matrix: $D^{-0.5}(D - A)\, D^{-0.5}$
  - *n x n* symmetric matrix



1.00	-0.52	-0.39	0.00	-0.06	0.00
-0.52	1.00	-0.50	0.00	0.00	0.00
-0.39	-0.50	1.00	-0.12	0.00	0.00
0.00	0.00	-0.12	1.00	0.47-	0.44-
-0.06	0.00	0.00	-0.47	1.00	0.50-
0.00	0.00	0.00	0.44-	0.50-	1.00

- Important properties:
  - Eigenvectors are real and normalize
  - Each A_{ij} (which i,j is not equal) = $-A_{ij} / D_{ii}$

An application: finding an optimal min-cut

- Express a bi-partition (A,B) as a vector

$$p_i = \begin{cases} +1 \ \text{if } x_i \in \mathrm{A} \\ -1 \ \text{if } x_i \in \mathrm{B} \end{cases} = p^T L\, p$$

Laplacian matrix

- The Laplacian is semi positive
- The Rayleigh Theorem shows:
  - The minimum value for f(p) is given by the <u>second smallest eigenvalue of the Laplacian</u>
  - The optimal solution for p is given by the eigenvector corresponding to the second smallest eigenvalue of the Laplacian, referred as the <u>Fiedler Vector</u>

Singular Value Decomposition

- If **A** is a rectangular m × k matrix of real numbers, then there exists an m × m orthogonal matrix **U** and a k × k orthogonal matrix **V** such that

$$\underset{(m\times k)}{\mathbf{A}} = \underset{(m\times m)}{\mathbf{U}}\ \underset{(m\times k)}{\Lambda}\ \underset{(k\times k)}{\mathbf{V}^T} \qquad \mathbf{U}\mathbf{U}^T = \mathbf{V}\mathbf{V}^T = \mathbf{I}$$

  - Λ is an m × k matrix where the (i, j)th entry λ_i ¸ 0, i = 1 $\cdots$ min(m, k) and the other entries are zero
    - The positive constants λ_i are the singular values of **A**
- If A has rank r, then there exists r positive constants λ_1, λ_2,$\cdots\lambda_r$, r orthogonal m × 1 unit vectors $\mathbf{u}_1,\mathbf{u}_2,\cdots,\mathbf{u}_r$ and r orthogonal k × 1 unit vectors $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_r$ such that

$$\mathbf{A} = \sum_{i=1}^{r} \lambda_i \mathbf{u}_i \mathbf{v}_i^T$$

  - Similar to the spectral decomposition theorem

SVD

- If **A** is a symmetric and positive definite then
  - SVD = Eigen decomposition
    - EIG(λ_i) = SVD(λ_i^2)
- Here $\mathbf{AA}^T$ has an eigenvalue-eigenvector pair (λ_i^2,u_i)

$$\begin{aligned}\mathbf{AA}^T &= \left(\mathbf{U}\Lambda\mathbf{V}^T\right)\left(\mathbf{U}\Lambda\mathbf{V}^T\right)^T \\ &= \mathbf{U}\Lambda\mathbf{V}^T\mathbf{V}\Lambda\mathbf{U}^T \\ &= \mathbf{U}\Lambda^2\mathbf{U}^T\end{aligned}$$

- Alternatively, the $\mathbf{v}_i$ are the eigenvectors of $\mathbf{A}^T\mathbf{A}$ with the same non zero eigenvalue λ_i^2

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\Lambda^2\mathbf{V}^T$$

Example for SVD

- Let **A** be a symmetric, positive definite matrix
  - **U** can be computed as

$$\mathbf{A}=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\Rightarrow \mathbf{A}\mathbf{A}^T=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\begin{bmatrix}3 & -1\\ 1 & 3\\ 1 & 1\end{bmatrix}=\begin{bmatrix}11 & 1\\ 1 & 11\end{bmatrix}$$

$$\det\left(\mathbf{A}\mathbf{A}^T-\gamma\mathbf{I}\right)=0\Rightarrow \gamma_1=12,\gamma_2=10\Rightarrow \mathbf{u}_1^T=\left[{1}/{\sqrt{2}},{1}/{\sqrt{2}}\right],\mathbf{u}_2^T=\left[{1}/{\sqrt{2}},{-1}/{\sqrt{2}}\right]$$

  - **V** can be computed as

$$\mathbf{A}=\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}\Rightarrow \mathbf{A}^T\mathbf{A}=\begin{bmatrix}3 & -1\\ 1 & 3\\ 1 & 1\end{bmatrix}\begin{bmatrix}3 & 1 & 1\\ -1 & 3 & 1\end{bmatrix}=\begin{bmatrix}10 & 0 & 2\\ 0 & 10 & 4\\ 2 & 4 & 2\end{bmatrix}$$

$$\det\left(\mathbf{A}^T\mathbf{A}-\gamma\mathbf{I}\right)=0\Rightarrow \gamma_1=12,\gamma_2=10,\gamma_3=0$$

$$\Rightarrow \mathbf{v}_1^T=\left[{1}/{\sqrt{6}},{2}/{\sqrt{6}},{1}/{\sqrt{6}}\right],\mathbf{v}_2^T=\left[{2}/{\sqrt{5}},{-1}/{\sqrt{5}},0\right],\mathbf{v}_3^T=\left[{1}/{\sqrt{30}},{2}/{\sqrt{30}},{-5}/{\sqrt{30}}\right]$$

Example for SVD

- Taking $\lambda^2_1=12$ and $\lambda^2_2=10$, the singular value decomposition of **A** is

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$$

$$= \sqrt{12}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 1/\sqrt{6}, 2/\sqrt{6}, 1/\sqrt{6} \end{bmatrix} + \sqrt{10}\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}\begin{bmatrix} 2/\sqrt{5}, -1/\sqrt{5}, 0 \end{bmatrix}$$

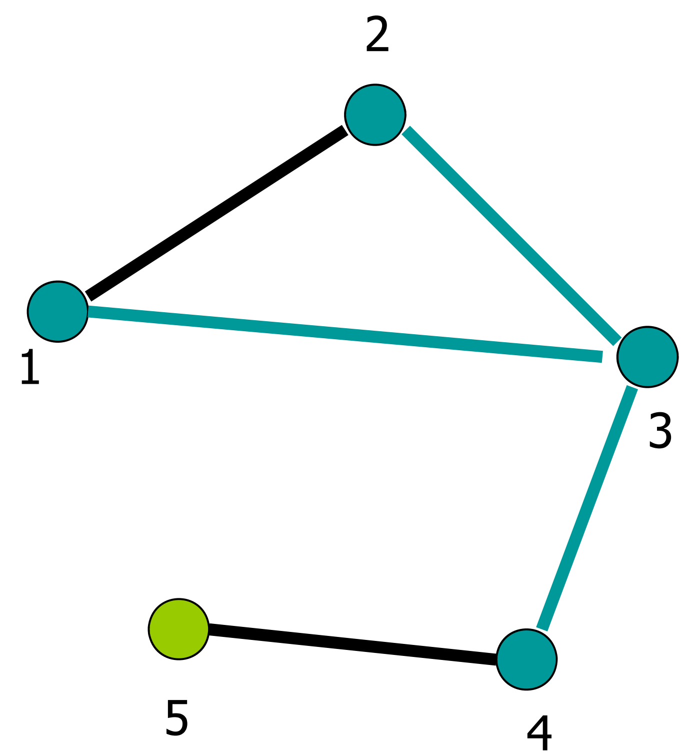
- Thus the U, V and Λ are computed by performing eigen decomposition of $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$
- Any matrix has a singular value decomposition but only symmetric, positive definite matrices have an eigen decomposition

Applications of SVD

- Inverse of an n × n square matrix, **A**
  - If **A** is non-singular, then $\mathbf{A}^{-1} = (\mathbf{U}\Lambda\mathbf{V}^T)^{-1} = \mathbf{V}\Lambda^{-1}\mathbf{U}^T$ where $\Lambda^{-1}=\mathrm{diag}(1/\lambda_1, 1/\lambda_1,\cdots, 1/\lambda_n)$
  - If **A** is singular, then $\mathbf{A}^{-1} = (\mathbf{U}\Lambda\mathbf{V}^T)^{-1} = \mathbf{V}\Lambda_0^{-1}\mathbf{U}^T$ where $\Lambda_0^{-1}=\mathrm{diag}(1/\lambda_1, 1/\lambda_2,\cdots, 1/\lambda_i,0,0,\cdots,0)$
- Least squares solutions of a m × n system
  - **Ax**=**b** (**A** is m£n, m¸n) =$(\mathbf{A}^T\mathbf{A})\mathbf{x}=\mathbf{A}^T\mathbf{b}$ **)** $\mathbf{x}=(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}=\mathbf{A}^+\mathbf{b}$
  - If $\mathbf{A}^T\mathbf{A}$ is singular, $\mathbf{x}=\mathbf{A}^+\mathbf{b}= (\mathbf{V}\Lambda_0^{-1}\mathbf{U}^T)\mathbf{b}$ where $\Lambda_0^{-1} = \mathrm{diag}(1/\lambda_1, 1/\lambda_2,\cdots, 1/\lambda_i,0,0,\cdots,0)$
- SVD can be used to compute optimal low-rank approximations of arbitrary matrices.

Subgraphs

- Subgraph: Given $V' \subseteq V$, and $E' \subseteq E$, the graph $G'=(V',E')$ is a subgraph of G.
- Induced subgraph: Given $V' \subseteq V$, let $E' \subseteq E$ is the set of all edges between the nodes in V'. The graph $G'=(V',E')$, is an induced subgraph of G



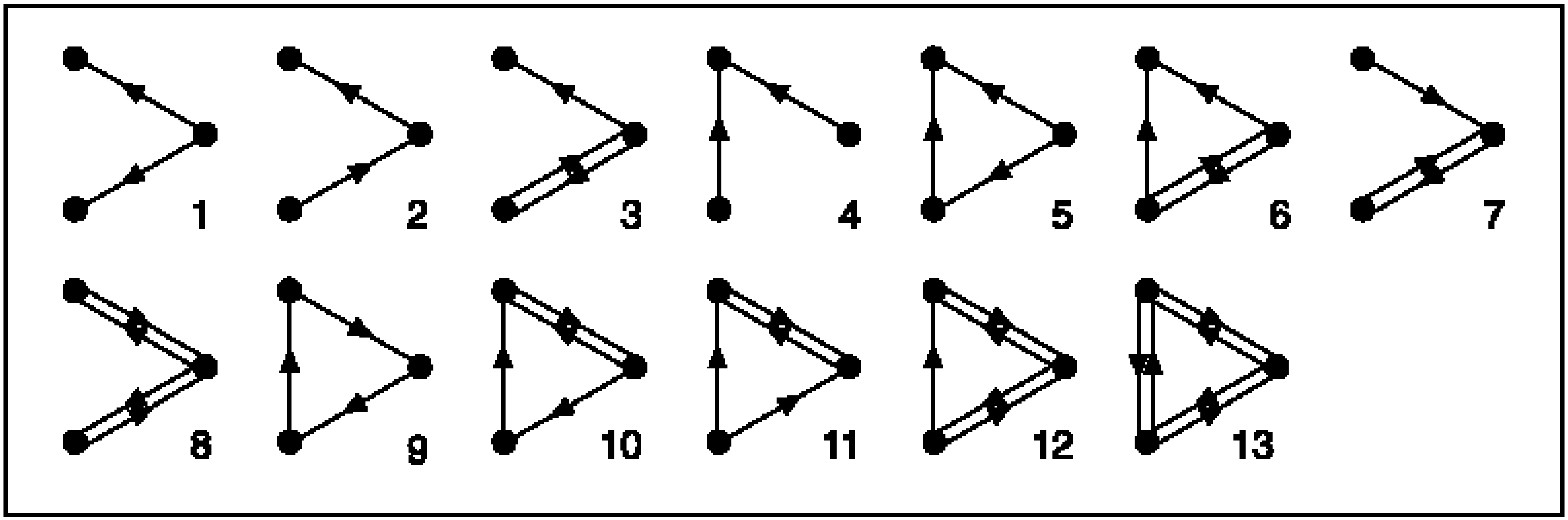
Motifs

- Some subgraphs might be important
  - Significantly overrepresented
  - Significantly underrepresented
- They are motifs
- Motifs as smallest building blocks of networks
- Various networks may have similar building blocks
- Their functionality may be related to their motif structure
  - Their synchronizability
  - Their cooperativity
  - There stability
  - Their robustness
- Motifs are important especially in biological networks

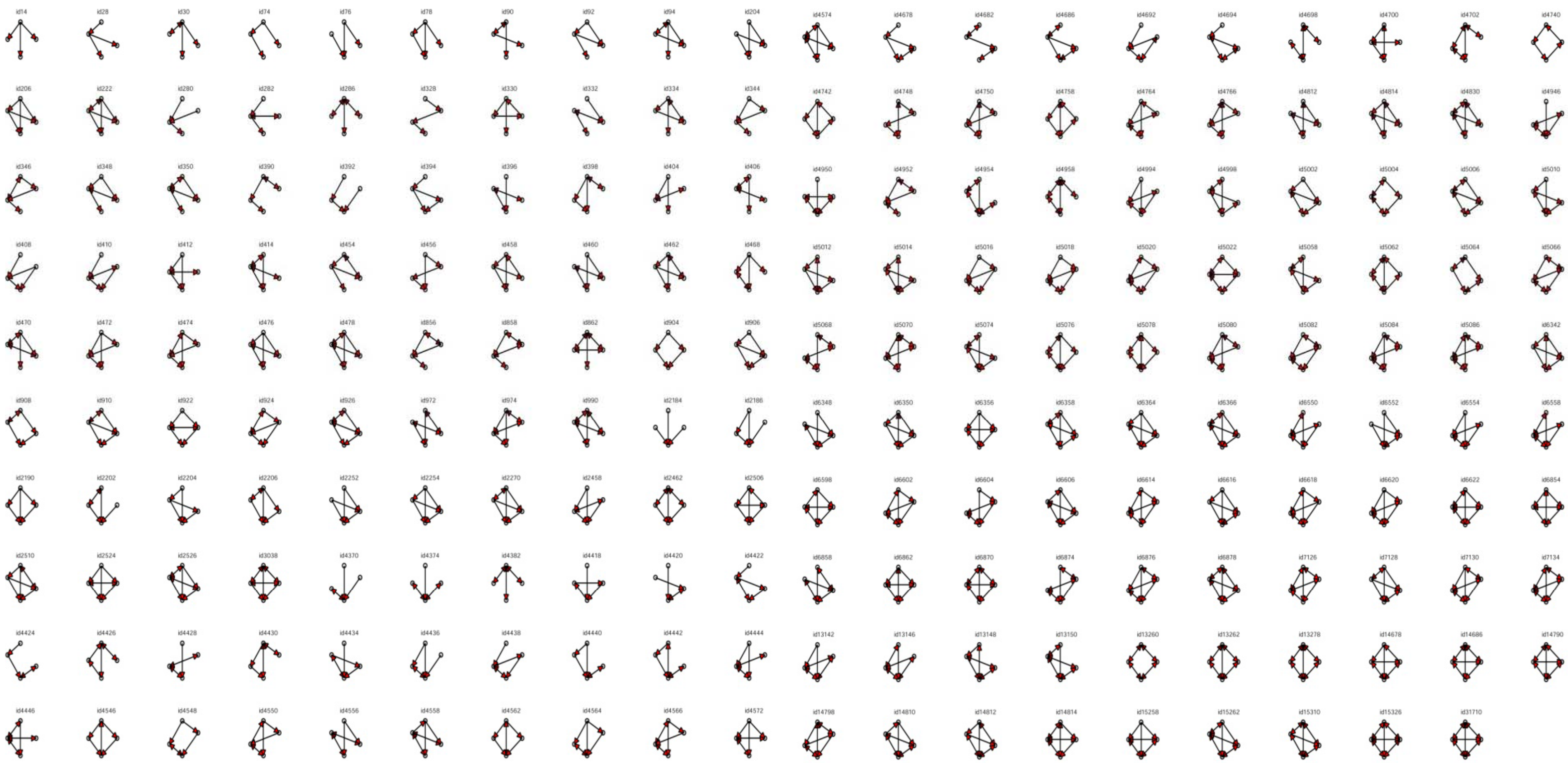
Motifs

- So, the challenge is to find the motifs!
- Can we define and detect building blocks of networks?
- What about comparing them with random networks!
  - Considering a subgraph
  - Counting its abundance in the network
  - Taking into account a number of proper random networks
  - Counting the abundance of the subgraph in random networks
  - Comparing them → the significance can be obtained
  - The subgraphs with significant difference in their appearance in the original network and random ones are called MOTIFS
- Done !

13 3-node connected subgraphs



199 4-node connected subgraphs



And it grows fast for larger subgraphs: **9364** 5-node subgraphs, **1,530,843** 6-node…

Motif detection process

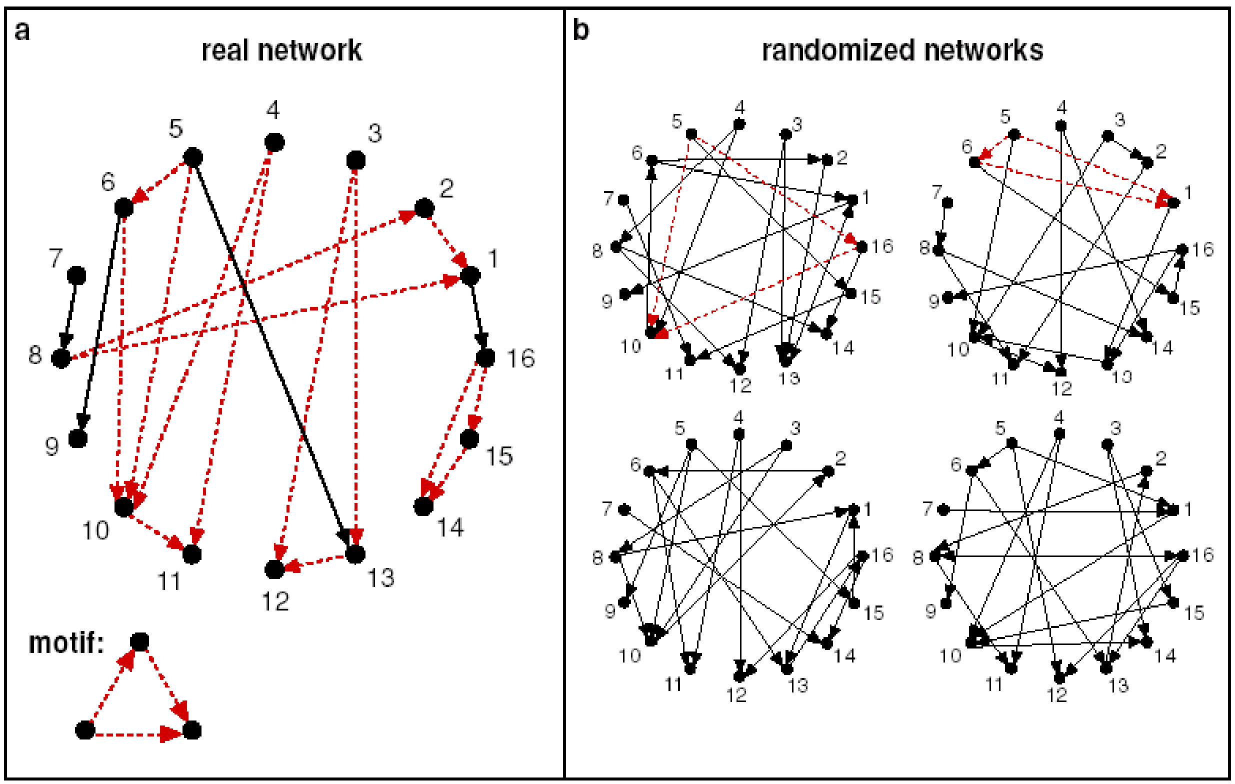
- Generation of a suitable random ensemble (reference networks): with the same in- and out-degree as the original network
- Count how many times each subgraph appears in the original network (N_{real})
- Count the statistics of the appearance of the subgraph in the random networks (with mean $\langle N_{rand} \rangle$ and standard deviation σ_{rand})
- Compute statistical significance – probability of appearing in random as much as in real network (P-val or Z-score)

$$Z = \frac{N_{real} - \langle N_{rand} \rangle}{\sigma_{rand}}$$

- Subgraphs with high Z-scores are denoted as **Network Motifs**

Motif detection process

- The idea: patterns that occur in the real network much more then in a randomized network, must have functional significance
- The randomized networks share the same number of edges and number of nodes, but edges are assigned at random

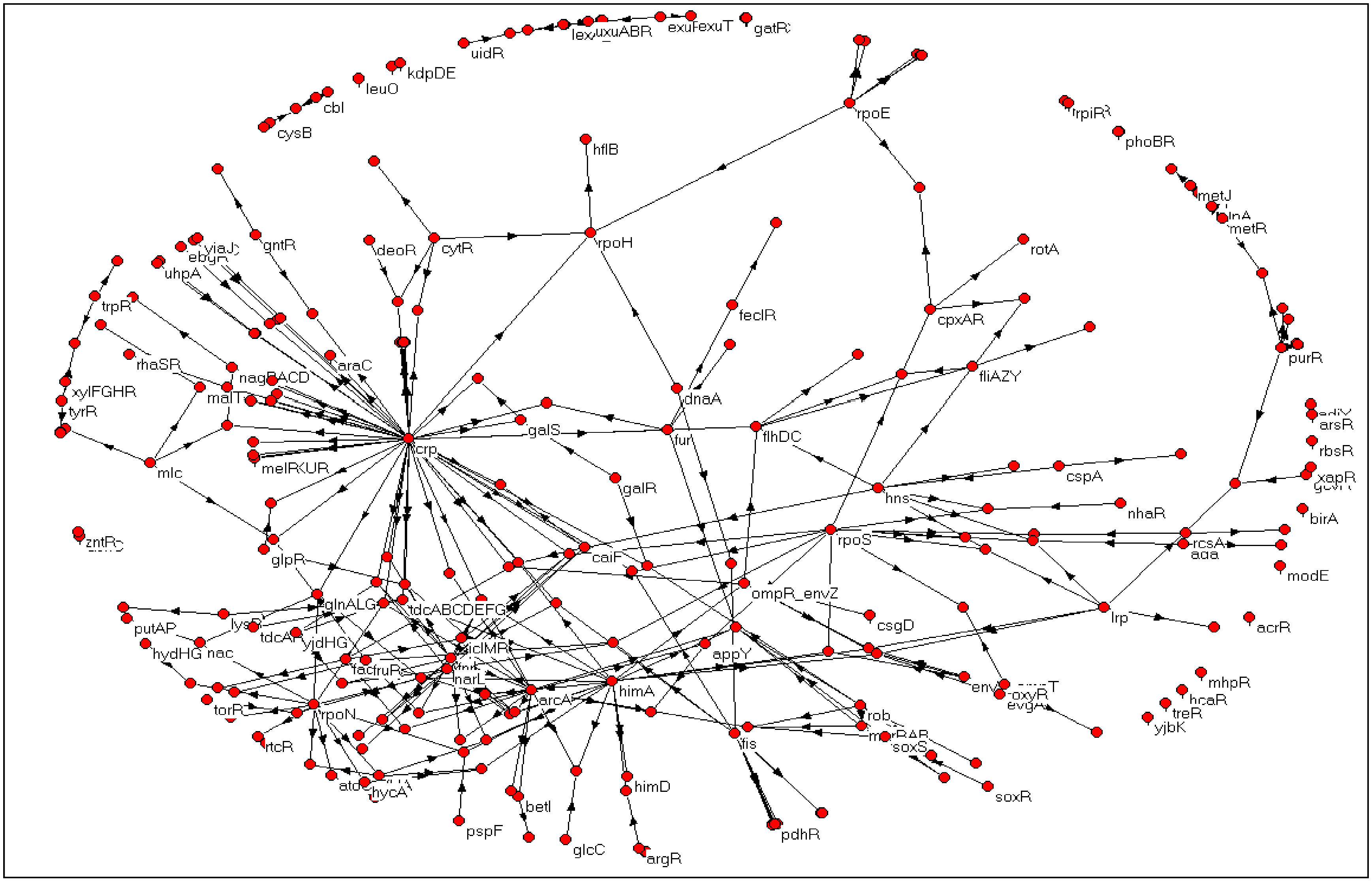


$N_{real} = 5$

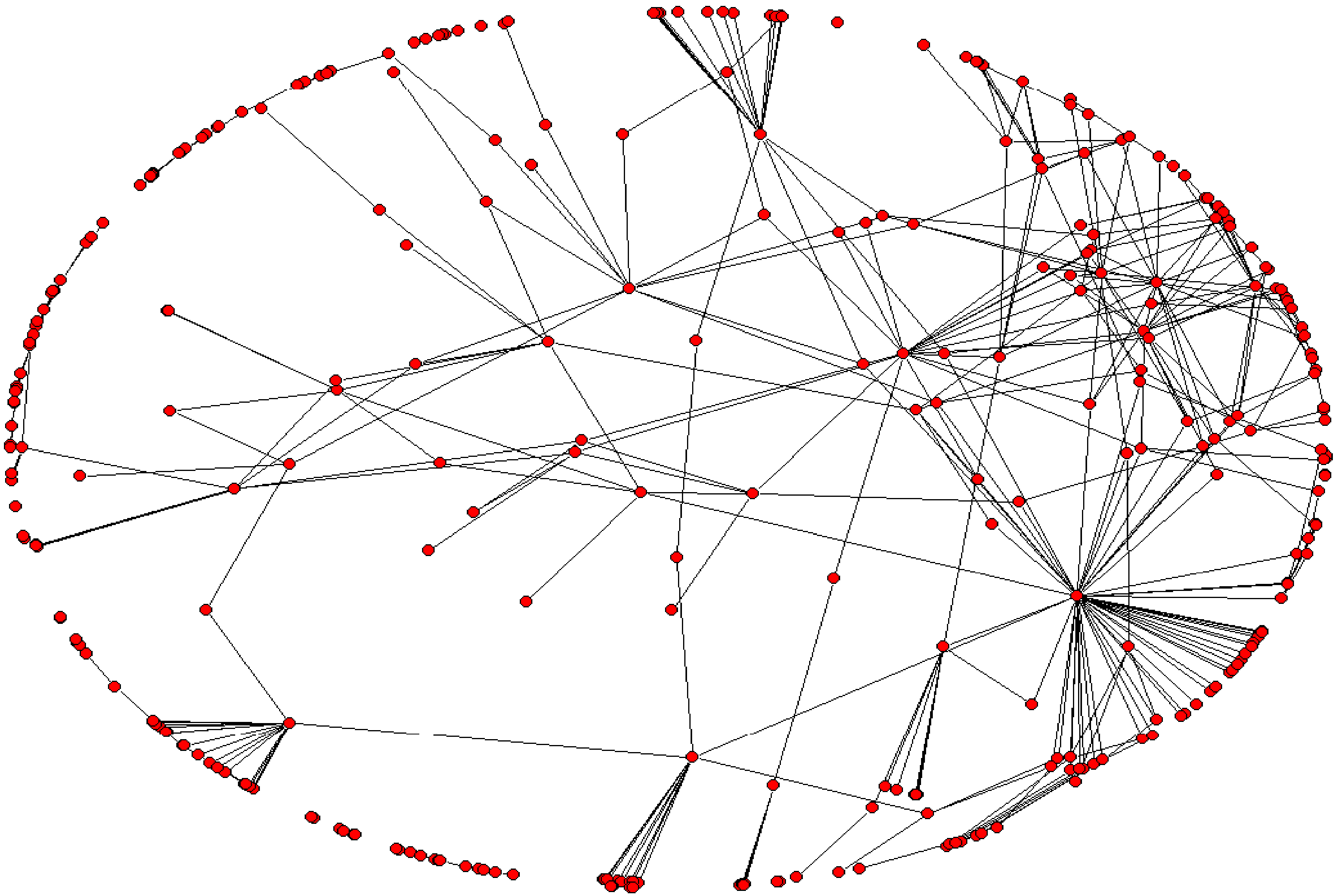
$N_{rand} = 0.5 \pm 0.6$

Z-score = 7.5

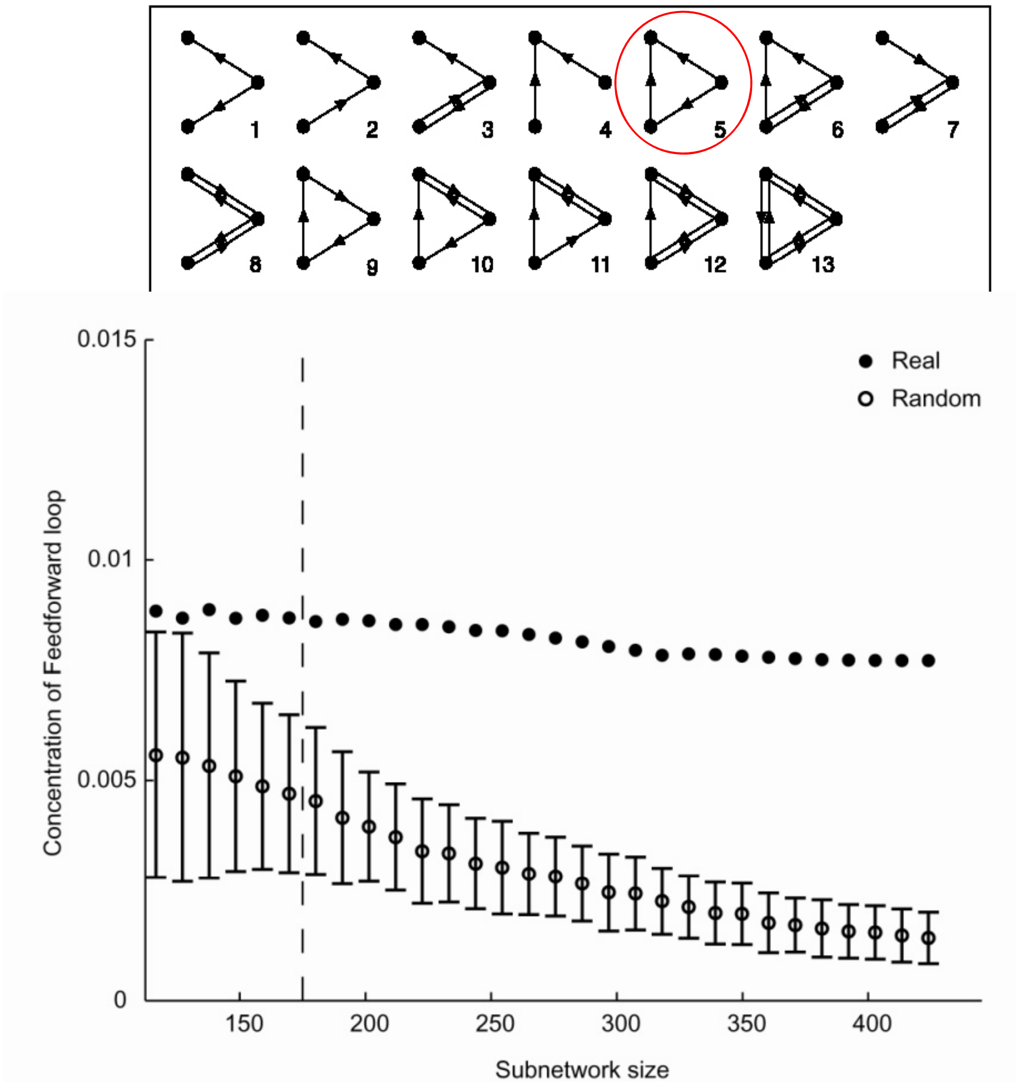
E. Coli transcription network

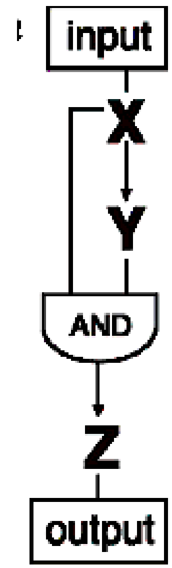


Randomized version of E. Coli



A 3-node motif





This motif is called feed-forward loop

N_{real} = 40

N_{rand} = 7±3

Z-Score = 10

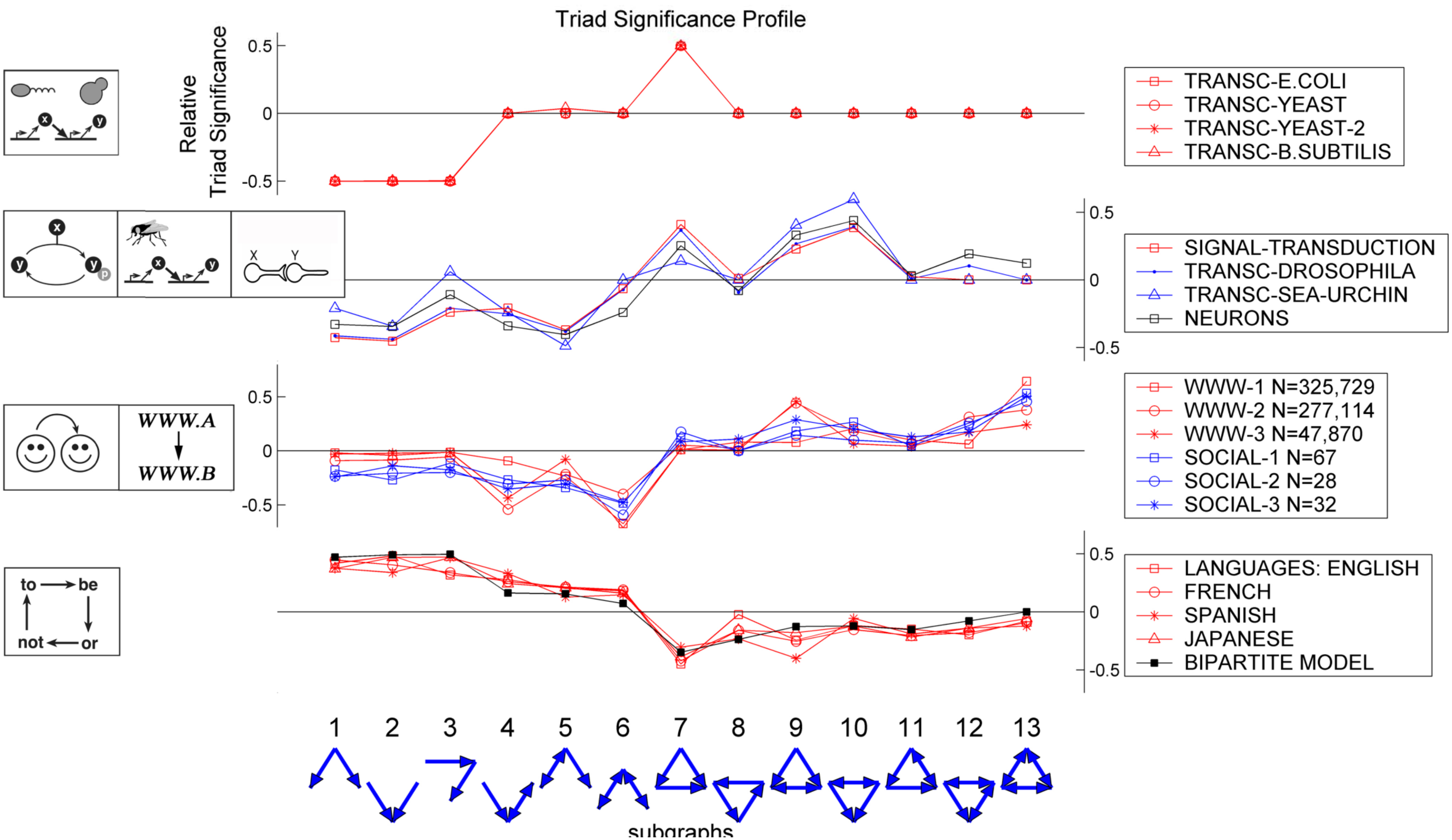
A similarity measure for motifs

- Various networks may have similar motif structures
- So, they might have been evolved with similar building blocks
- How can we compare networks of different sizes for similarity in local structure?
- Subgraph significance profile goes beyond detection of motifs
- Definition: the set of normalized z-scores of all subgraphs of given size

$$SSP_i = \frac{Z_i}{\sqrt{\sum Z_i^{\ 2}}}$$

- Advantages:
  - Highlights the relative importance of motifs
  - Identifies anti-motifs

3-node motif significance profile



Isomorphism

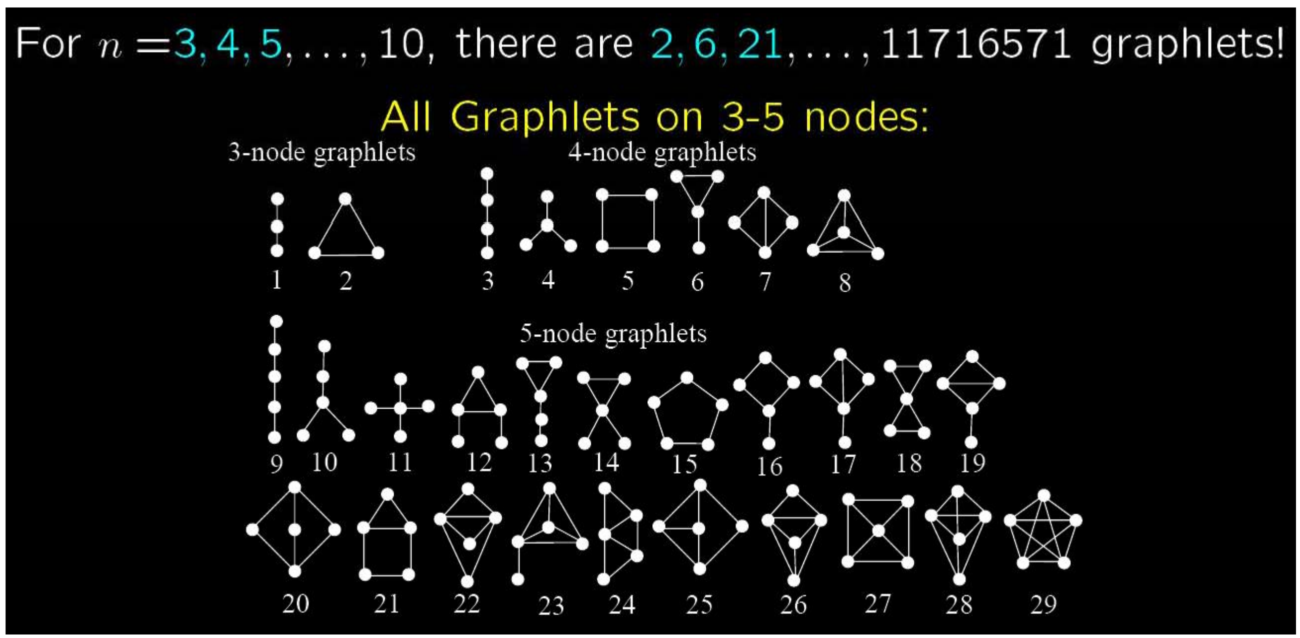
- In graph theory, an **isomorphism of graphs** G and H is a bijection between the vertex sets of G and H

$$f: V(G) \to V(H)$$

- such that any two vertices *u* and *v* of *G* are adjacent in *G* ***if and only if*** $f(u)$ and $f(v)$ are adjacent in *H*.
- This kind of bijection is commonly called "edge-preserving bijection", in accordance with the general notion of isomorphism being a structure-preserving bijection.

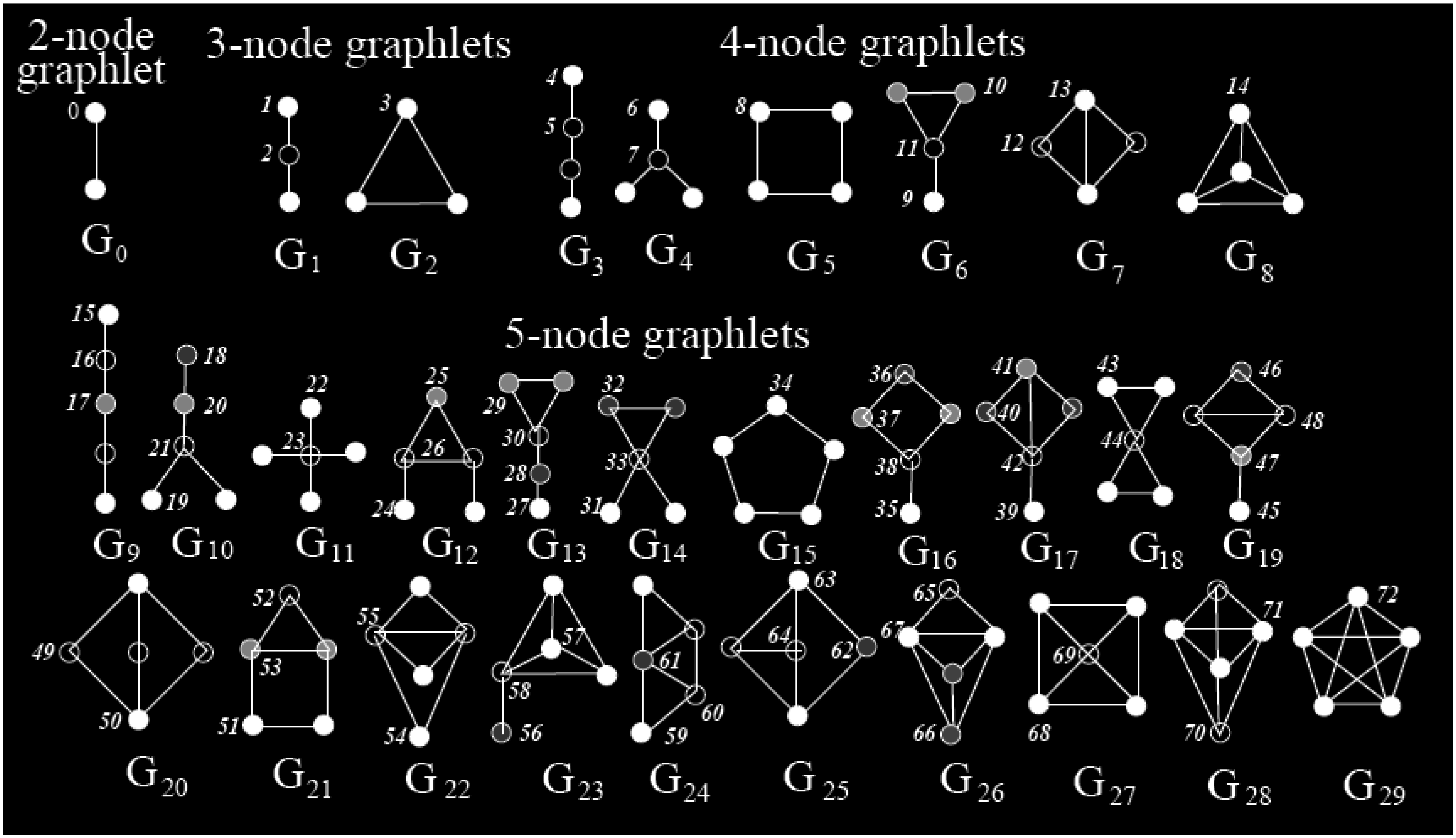
Graphlets

- The motif definition is sensitive to the randomization method in producing the random networks
- An Alternative is to consider GRAPHLETS
- Definition: graphlets are small connected non-isomorphic subgraphs of a graph G induced on $n \geq 3$ nodes of G
- Different from network motifs
  - Induced subgraphs
  - Of any frequency



N. Przulj, D. G. Corneil, and I. Jurisica, "Modeling Interactome: Scale Free or Geometric?," *Bioinformatics*, vol. 20, num. 18, pg. 3508-3515, 2004.

Graphlets



2-node graphlet
3-node graphlets
4-node graphlets
5-node graphlets
G_0 G_1 G_2 G_3 G_4 G_5 G_6 G_7 G_8
G_9 G_{10} G_{11} G_{12} G_{13} G_{14} G_{15} G_{16} G_{17} G_{18} G_{19}
G_{20} G_{21} G_{22} G_{23} G_{24} G_{25} G_{26} G_{27} G_{28} G_{29}

Readings

- Leslie Hogben, Spectral Graph Theory and the Inverse Eigenvalue Problem of a Graph; Chamchuri Journal of Mathematics, Volume 1(2009) Number 1, 51–72
- Milo et al, Network Motifs: Simple Building Blocks of Complex Networks, Science 2002; Superfamilies of Evolved and Designed Networks, Science 2004