Σύνθετα Δίκτυα

com+plex: with+ -fold (having parts)

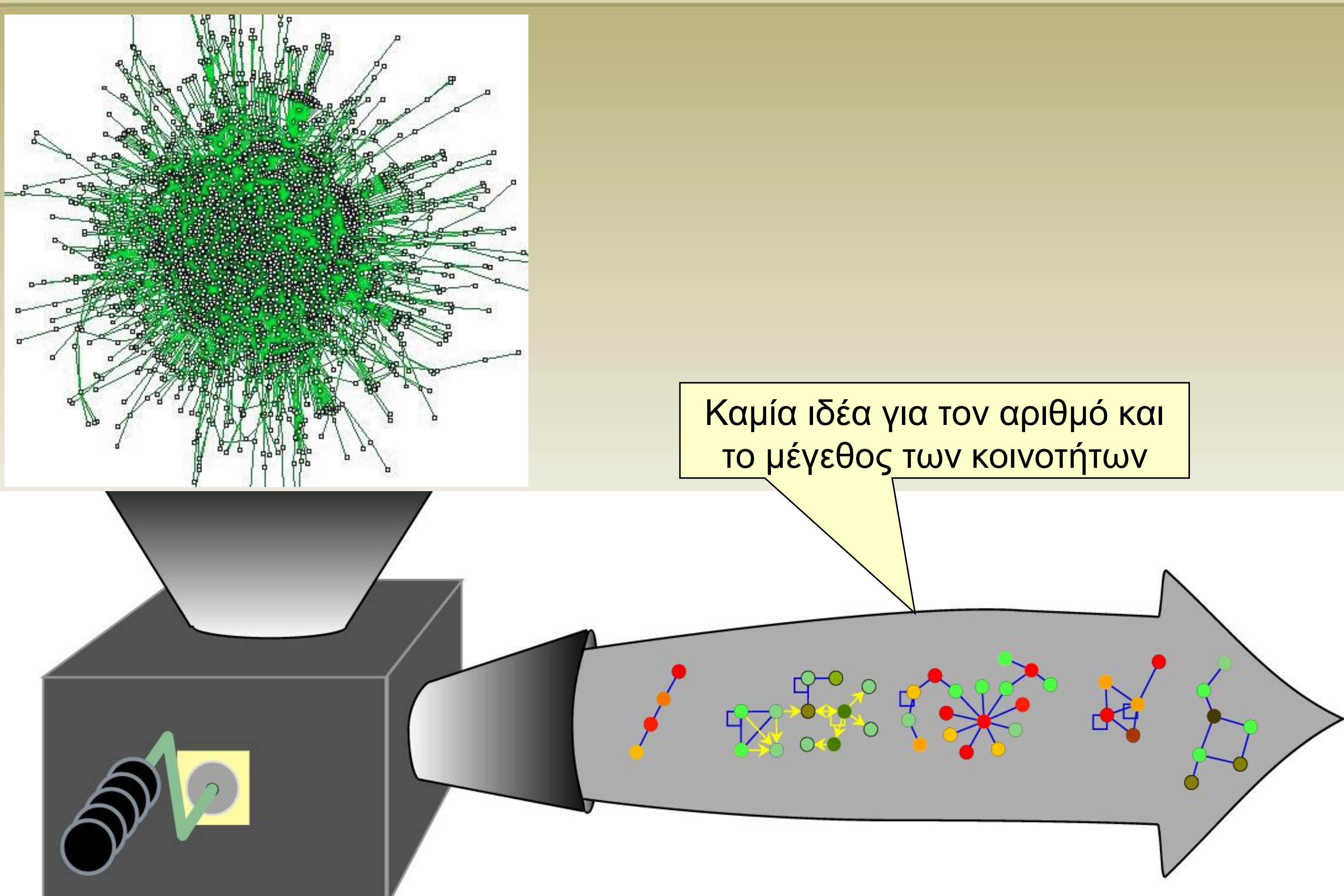
Διδάσκων –
Δημήτριος Κατσαρός

Διάλεξη 11η

Communities in Complex Networks

Κοινότητες σε Σύνθετα Δίκτυα

Διυλίζοντας δίκτυα για κοινότητες



Modularity (Q) μιας διαίρεσης

Q = #(ακμών εντός groups) - E(#(ακμών εντός groups σε ένα RANDOM γράφημα με τους ίδιους βαθμούς κόμβων))

Trivial division: όλοι οι κόμβοι σε ένα group

==> Q(trivial division) = 0

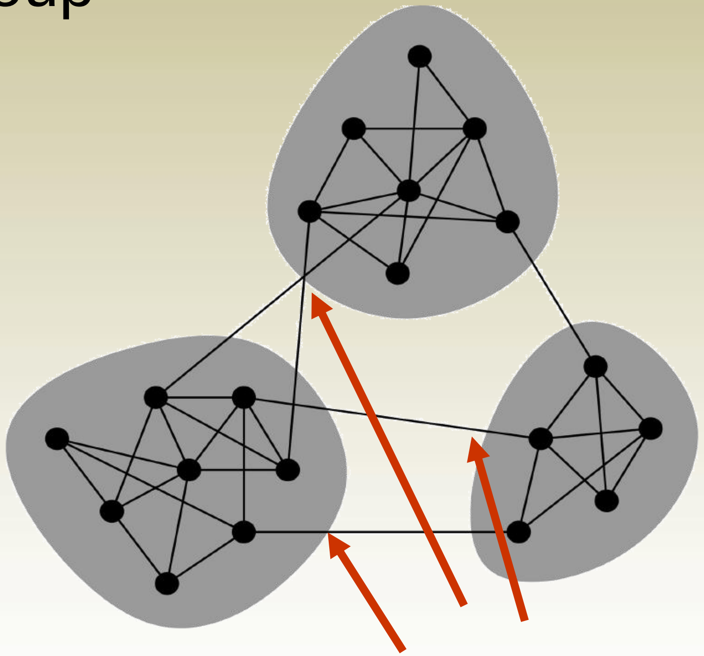
k_i = βαθμός κόμβου i

$M = \sum k_i = 2|E|$

A_{ij} = 1 εάν $(i,j) \in E$, 0 αλλιώς

E_{ij} = αναμενόμενος αριθμός ακμών μεταξύ i και j σε ένα random γράφημα με τους ίδιους βαθμούς.

Λήμμα: $E_{ij} \approx k_i * k_j / M$



Ακμές μεταξύ groups

$$Q = \sum(A_{ij} - k_i * k_j / M \mid i,j \text{ στο ίδιο group})$$

Αλγόριθμος 1: Διαίρεση σε δυο groups

$Q = \sum(A_{ij} - k_i*k_j/M \ \ | \text{ i,j στο ίδιο group})$

- Έστω ότι έχουμε n κόμβους {1,…,n}
- **s** - {±1} διάνυσμα μεγέθους n
  Αναπαριστούμε μια 2-διαίρεση:
  - $s_i == s_j$ iff i και j βρίσκονται στο ίδιο group
  - ½ $(s_i*s_j+1) = 1$ εάν $s_i==s_j$, 0 αλλιώς
- ==> $$\boxed{Q = \frac{1}{2}\sum_{i,j}(A_{ij} - \frac{k_i k_j}{M})(s_i s_j + 1)}$$

Αλγόριθμος 1: Διαίρεση σε δυο groups

$$Q = \frac{1}{2}\sum_{i,j}(A_{ij} - \frac{k_i k_j}{M})(s_i s_j + 1)$$

Αφού $\sum_{i,j} A_{ij} = \sum_i k_i = M$

$$Q = \frac{1}{2}\sum_{i,j}(A_{ij} - \frac{k_i k_j}{M})s_i s_j$$

B = ο <u>modularity</u> πίνακας

- συμμετρικός
- row sum = 0

0 είναι μια eigvenvalue του **B**

$$Q = \frac{1}{2}\boldsymbol{s}^T \boldsymbol{B} \boldsymbol{s}$$ όπου $$B_{ij} = A_{ij} - \frac{k_i k_j}{M}$$

Πίνακας modularity : Παράδειγμα

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

$$k_1 = k_2 = 1,$$

$$k_3 = k_4 = k_5 = 2$$

$$M = 8$$

$$\boxed{B_{ij} = A_{ij} - \frac{k_i k_j}{M}}$$

$$\boldsymbol{B} = \frac{1}{8} \begin{pmatrix} -1 & 7 & -2 & -2 & -2 \\ 7 & -1 & -2 & -2 & -2 \\ -2 & -2 & -4 & 4 & 4 \\ -2 & -2 & 4 & -4 & 4 \\ -2 & -2 & 4 & 4 & -4 \end{pmatrix}$$

Αλγόριθμος 1: Διαίρεση σε δυο groups

B συμμετρικός ⇒ **B** διαγωνιοποιήσιμος (real eigenvalues)

B eigenvalues

$$\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n$$

B αντίστοιχα eigenvectors

$$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots \boldsymbol{u}_n$$

$$\boldsymbol{B}\boldsymbol{u}_i = \beta_i \boldsymbol{u}_i$$

$$\boxed{Q = \frac{1}{2}\boldsymbol{s}^T \boldsymbol{B} \boldsymbol{s}} \Rightarrow \boxed{Q = \frac{1}{2}\sum_i \beta_i a_i^2}$$

$$\text{where } a_i = \boldsymbol{u}_i^T \cdot \boldsymbol{s}$$

$$\left(\boldsymbol{s} = \sum_i a_i \boldsymbol{u}_i \right) \quad n = \|\boldsymbol{s}\|^2 = \sum a_i^2$$

- Ποια διανύσματα μεγιστοποιούν το Q?
  - προφανώς **s** ~ $\mathbf{u}_1$ μεγιστοποιεί το Q, αλλά $\mathbf{u}_1$ μπορεί να μην είναι {±1} διάνυσμα
  - Άπληστο ευριστικό: επιλογή **s** ~ **u1**: s_i= +1 if u_i>0, s_i=-1 αλλιώς

Algorithm 1 Dividing into two modules

Input: A network with n vertices, $n > 1$

1: Compute the leading eigenpair, $\boldsymbol{u}_1$ and β_1, of the modularity matrix $\boldsymbol{B}$.
2: **if** β_1 is positive **then**
3:     Compute the vector $\boldsymbol{s} = (s_1, s_2, \ldots, s_n)$ with $s_i = +1$ if the corresponding element in $\boldsymbol{u}_1$ is positive, and $s_i = -1$ otherwise.
4:     **if** $\boldsymbol{s}^T \boldsymbol{B} \boldsymbol{s}$ is positive **then**
5:         **return** a division of the vertices into two groups according to the signs of the elements in $\boldsymbol{s}$.
6:     **else**
7:         **return** the network is indivisible
8:     **end if**
9: **else**
10:     **return** the network is indivisible
11: **end if**

$$\boxed{Q = \frac{1}{2} \boldsymbol{s}^T \boldsymbol{B} \boldsymbol{s}}$$

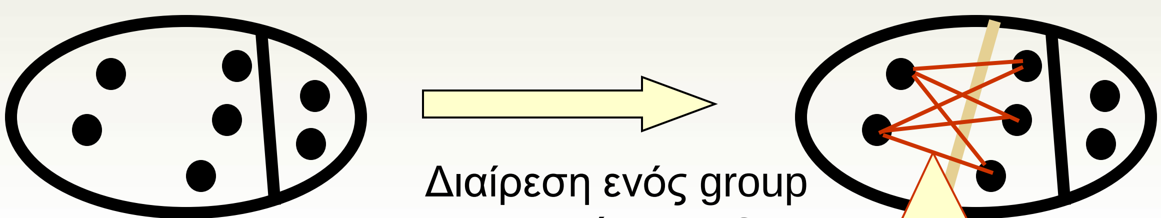
$$\boxed{Q = \frac{1}{2} \sum_i \beta_i a_i^2}$$

Διαίρεση σε περισσότερα από δυο groups

- Πώς να κάνουμε διαίρεση σε περισσότερα από δυο groups?
- <u>Εφαρμογή του αλγορίθμου αναδρομικά σε κάθε group</u>

$$\boxed{Q = \frac{1}{2}\sum_{i,j}\underbrace{(A_{ij} - \frac{k_i k_j}{M})}_{\text{Bij}}\underbrace{(s_i s_j + 1)}_{0|1}}$$

=1 iff i και j είναι στο ίδιο group, 0 αλλιώς



{i,j} ζεύγη που πρέπει να ενημερωθούν στο Q

Διαίρεση σε περισσότερα από δυο groups

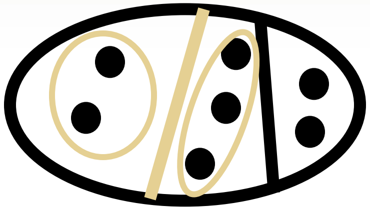
$$Q = \frac{1}{2}\sum_{i,j}\underbrace{(A_{ij} - \frac{k_i k_j}{M})}_{\text{Bij}}\underbrace{(s_i s_j + 1)}_{0|1}$$

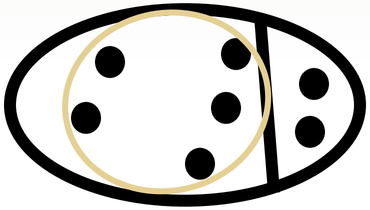
- g – ένα group με n_g κόμβους
- **s** – ένα {±1} διάνυσμα μεγέθους n_g
- Υπολογισμός του ΔQ για μια 2-διαίρεση του g

$$\Delta Q \;=\; \frac{1}{2}\sum_{i,j\in g} B_{ij}(s_i s_j + 1) - \sum_{i,j\in g} B_{ij}$$

Νέο: Τα στοιχεία του g σπάζουν σε δυο υπο-ομάδες (corresponding to s)

Παλιό: όλα τα σοιχεία του g ανήκουν στην ίδια ομάδα (g)





Διαίρεση σε περισσότερα από δυο groups

$$\Delta Q = \frac{1}{2}\sum_{i,j\in g} B_{ij}(s_i s_j + 1) - \sum_{i,j\in g} B_{ij}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\Delta Q = \frac{1}{2}\sum_{i,j\in g}\Big(B_{ij} - \delta_{ij}\underbrace{\sum_{k\in g} B_{ik}}_{f_i^{(g)}}\Big)s_i s_j$$

B[g] = ο υποπίνακας του B ορισμένος από g

$$\Delta Q = \frac{1}{2}\boldsymbol{s}^T\widehat{\boldsymbol{B}}[g]\boldsymbol{s}$$

όπου

$$\widehat{B}[g]_{ij} = B[g]_{ij} - \delta_{ij} f_i^{(g)}$$

γενικευμένος πίνακας modularity

$f_i^{(g)}$ = sum της i-οστής γραμμής B[g]

$f_i^{(\{1,\ldots,n\})} = 0$

Γενικευμένος πίνακας modularity: Παράδειγμα

$$B = \frac{1}{8}\begin{pmatrix} -1 & 7 & -2 & -2 & -2 \\ 7 & -1 & -2 & -2 & -2 \\ -2 & -2 & -4 & 4 & 4 \\ -2 & -2 & 4 & -4 & 4 \\ -2 & -2 & 4 & 4 & -4 \end{pmatrix}$$

$$\boldsymbol{B}[g] = \frac{1}{8}\begin{pmatrix} -1 & -2 & -2 \\ -2 & -4 & 4 \\ -2 & 4 & -4 \end{pmatrix}$$

g = {1, 4, 5}
(1 είναι ο minimal index)

$$\boldsymbol{f}^{(g)} = \frac{1}{8}(-5, -2, -2)$$

$$\widehat{\boldsymbol{B}}[g] = \frac{1}{8}\begin{pmatrix} 4 & -2 & -2 \\ -2 & -2 & 4 \\ -2 & 4 & -2 \end{pmatrix}$$

Algorithm 2 Divide a network module into two

Input: A module g in a network

Compute the leading eigenpair, $\boldsymbol{u}_1$ and β_1, of the modularity matrix $\widehat{\boldsymbol{B}}[g]$.

if β_1 is positive **then**

  Compute the vector $\boldsymbol{s} = (s_1, s_2, \ldots, s_{n_g})$ with $s_i = +1$ if the corresponding element in $\boldsymbol{u}_1$ is positive, and $s_i = -1$ otherwise.

  / Optional improvement: Refine partition by calling Algorithm 4 (see Section 2.4) */*

  if $\boldsymbol{s}^T \widehat{\boldsymbol{B}}[g] \boldsymbol{s}$ is positive **then**

    Let g_1 be the set of positive indices in $\boldsymbol{s}$.

    return g_1

  end if

end if

return g

Algorithm 3 Dividing a network into modularity groups

Input: A network with n vertices, $n > 1$

$P \leftarrow \{\{1, 2, \ldots, n\}\}$ */* A trivial partition in which all the vertices belo*

Mark the single group in P as divisible

while there are divisible groups in P **do**

  $g \leftarrow$ a divisible group in P

  $g_1 \leftarrow$ the return value of Algorithm 2 on the module g

  if $|g_1| = 0$ **or** $|g_1| = |g|$ **then**

    Mark g as indivisible

  else

    $g_2 \leftarrow g \setminus g_1$

    $P = P \setminus \{g\} \cup \{g_1, g_2\}$

    for $i = 1$ to 2 **do**

      if $|g_i| = 1$ **then**

        Mark g_i as indivisible

      else

        Mark g_i as divisible

      end if

    end for

  end if

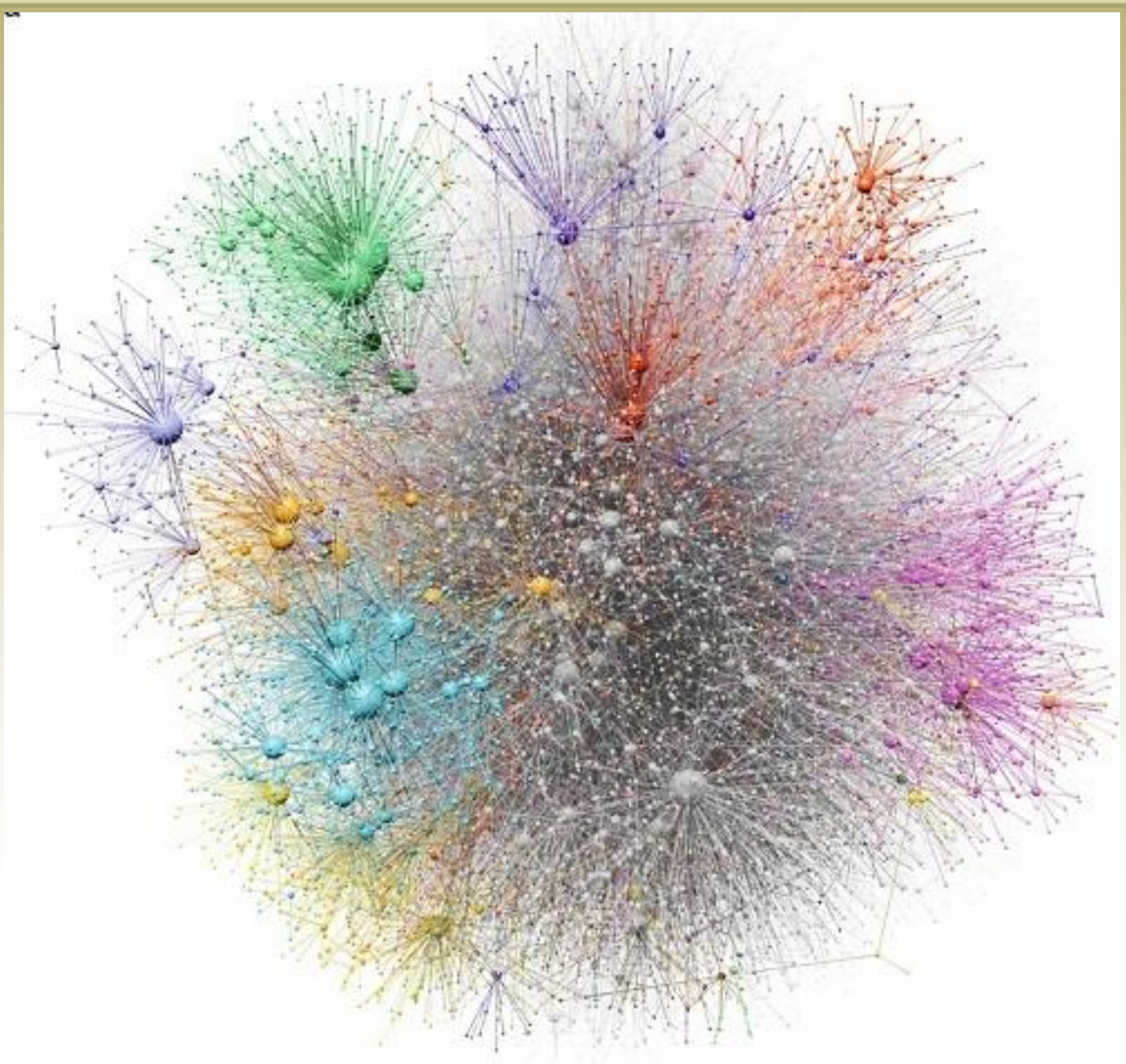
end while

Παράδειγμα του αλγορίθμου Louvain γαι την μεγιστοποίηση της modularity

Οπτικοποίηση του Internet (στο επίπεδο Autonomous Systems) με 22963 κόμβους και 48436 συνδέσεις

The layout generated by Graph Neural Network *NeuLay-2*

We used the *Louvain algorithm* to identify the community structure of the network and for visual clarity we highlight 12 communities in color



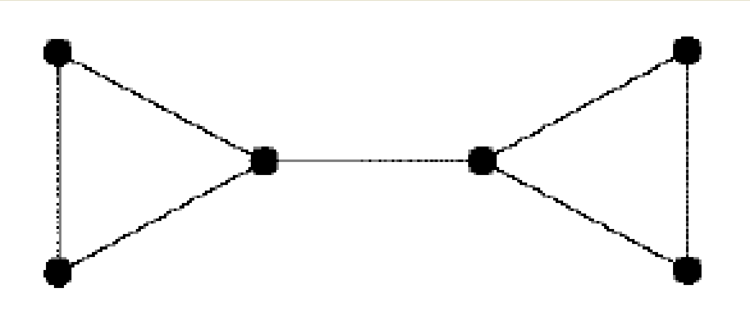
From the article:

https://www.nature.com/articles/s41467-023-37189-2

Ασκήσεις

- **Άσκηση 1.**

  Κατασκευάστε τον modularity matrix **B** για το ακόλουθο δίκτυο.



Βρείτε το ιδιοδιάνυσμα του **B** που αντιστοιχεί στην μεγαλύτερη ιδιοτιμή, και άρα διαμερίστε το δίκτυο σε δυο κοινότητες.