

METODE INTELIGENTE DE REZOLVARE A PROBLEMELOR REALE



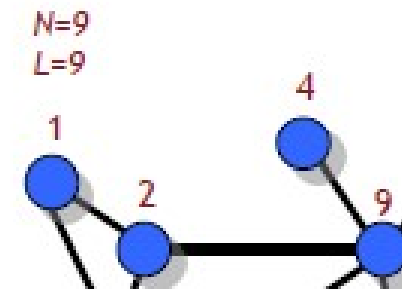
Laura Dioşan
Tema 5

Rețele complexe

- Ce este o rețea complexă?
- Concepte de bază
- Tipuri de rețele complexe
- Aplicații

Ce este o rețea complexă?

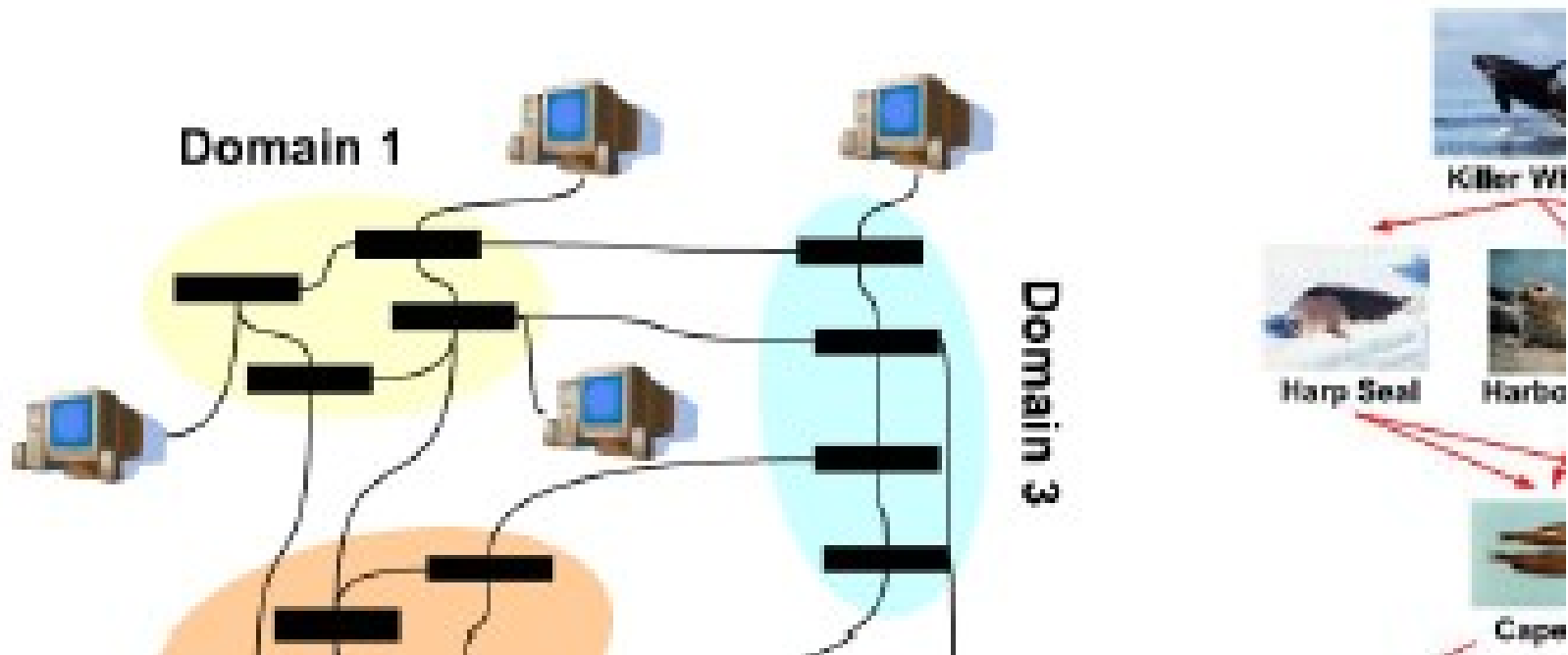
- O mulțime de elemente cu conexiuni între ele
- O rețea (un graf) $G=(N,M)$ este format dintr-o mulțime de noduri $N=\{n_1, n_2, \dots, n_N\}$ și o mulțime de conexiuni $L=\{l_1, l_2, \dots, l_M\}$
- Un graf = o sbractizare matematică a rețelei



- Fiecare element este reprezentat prin
 - Locație (fizică)
 - Nod (informatică)
 - Actor (sociologie)
 - Vârf (teoria grafelor)
- Interacțiunea dintre 2 elemente este reprezentată prin
 - Legătură (fizică)
 - Link (informatică)
 - Relație (sociologie)
 - Muchie (teoria grafelor)

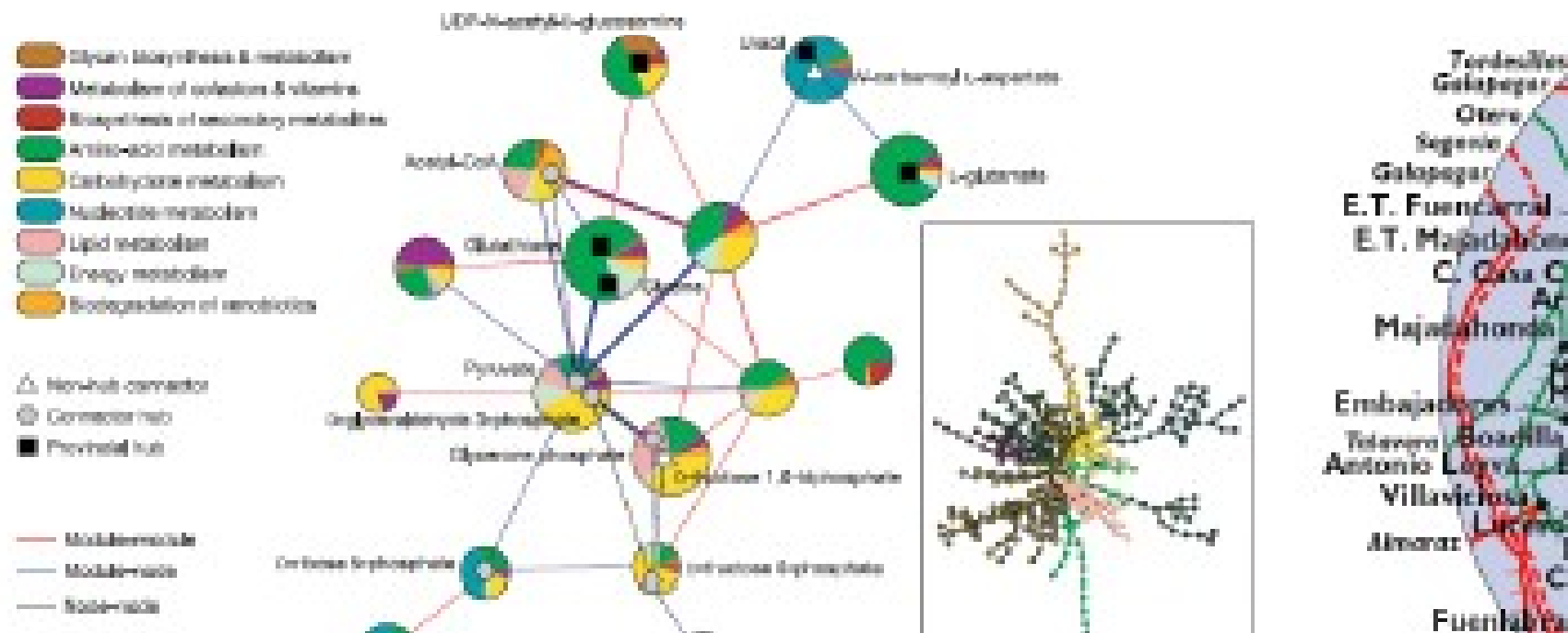
Ce este o rețea complexă?

- Nodurile și legăturile pot proveni din contexte variate



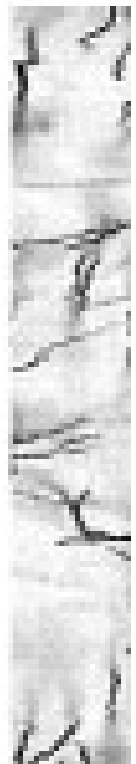
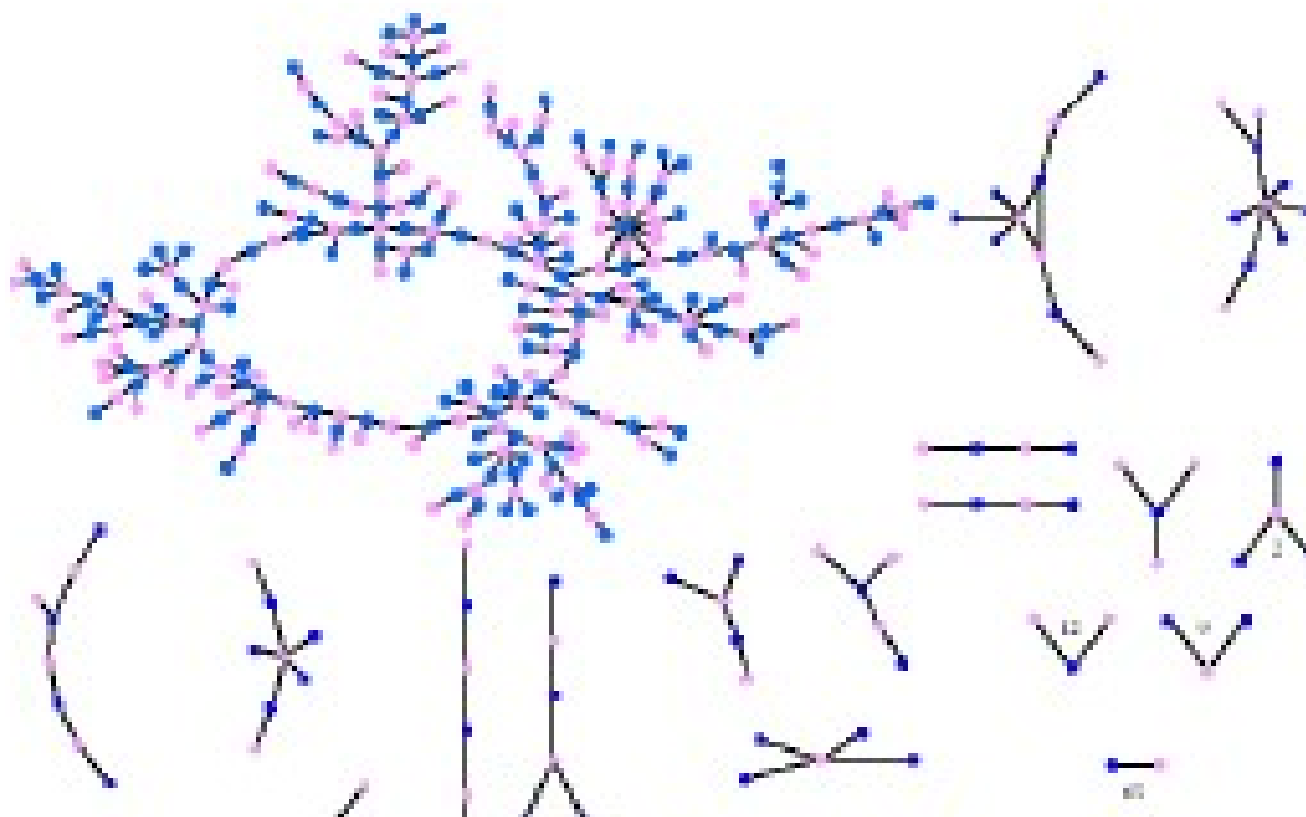
Ce este o rețea complexă?

- ❑ Nodurile și legăturile pot proveni din contexte variate



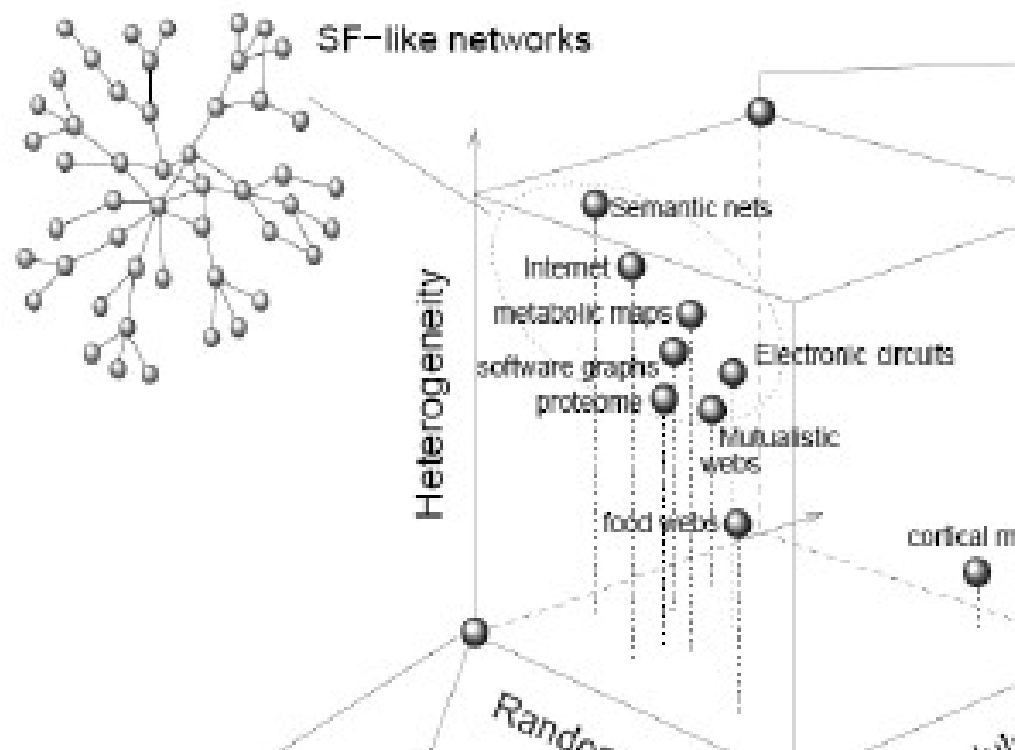
Ce este o rețea complexă?

- Nodurile și legăturile pot proveni din contexte variate



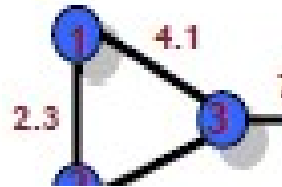
Ce este o rețea complexă?

- O rețea complexă este o rețea cu
 - attribute topologice ne-triviale,
 - patternuri de conexiuni care nu sunt nici pur aleatorii, dar nici pur regulate



Concepte de bază

- Descrierea rețelei (cu N noduri și M legături) prin matrici



- Matricea ponderilor fiecărei legături

0.0	2.3	4.1
2.3	0.0	1.0
4.1	1.0	0.0

- Matricea de adiacență A

0	1	1
1	0	1
1	1	0

- Matricea Laplaciană $L = A - K$

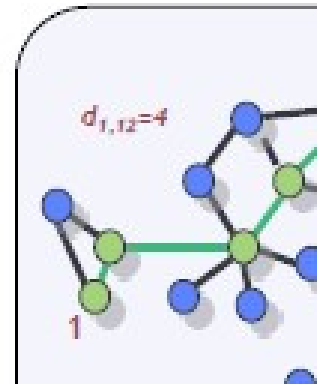
- K – matrice diagonală cu

$$k_{ii} = \sum a_{ij}$$

-2	1	1
1	-2	1
1	1	-3

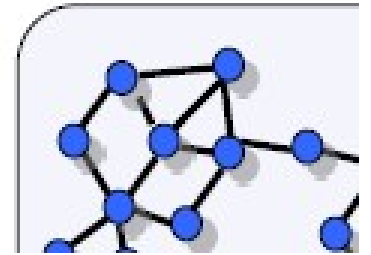
Concepte de bază

- Cel mai scurt drum între 2 noduri (d_{ij})
 - Cel mai scurt drum (ca nr de muchii)
 - Cel mai scurt drum (ca sumă a ponderilor muchiilor)
- Drum mediu /
 - Media (aritmetică sau armonică) tuturor celor mai scurte drumuri între oricare 2 noduri ale rețelei
- Diametru (D)
 - Cel mai lung dintre toate drumurile cele mai scurte
- Componentă
 - Mulțimea tuturor nodurilor în care se poate ajunge plecând dintr-un anumit nod



Concepte de bază

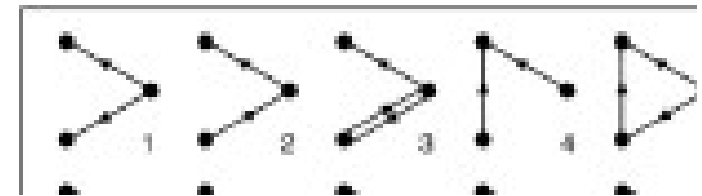
- ❑ Gradul unui nod (k_i)
 - Numărul de legături ale unui nod
 - ❑ Puterea unui nod (s_i)
 - Suma ponderilor conexiunilor unui nod $s_i = \sum w_{ij}$
 - ❑ Apropierea (b_i) unui nod sau a unei legături
 - Numărul celor mai scurte drumuri care trec prin acel nod/acea legătură
-
- ❑ Nodul cu cel mai mare grad e și cel mai puternic?
 - ❑ Care e nodul cu cea mai mare apropiere?



Concepte de bază

- Motive în rețea
 - Componente (sub-grafe) care apar în rețea mai des decât ne-am aștepta (în caz aleator)
 - Fiecare motiv poate "coda" informații specifice

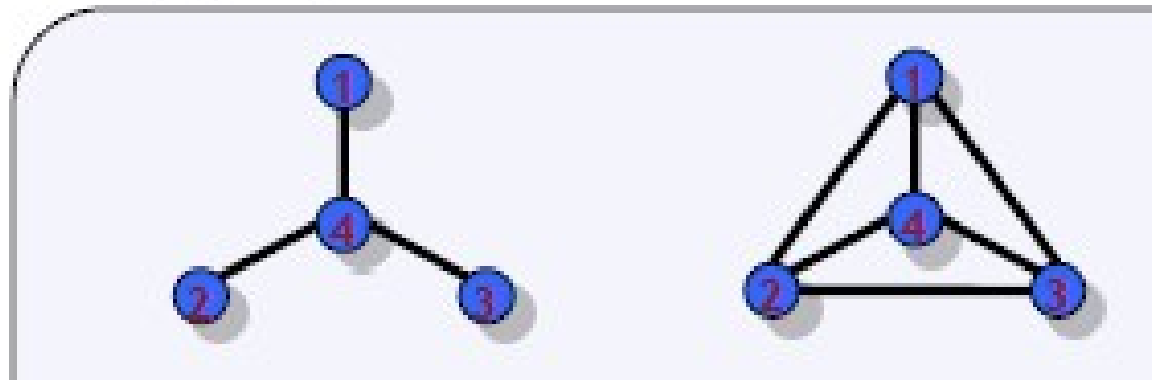
Example: all 13 types of three-n subgraphs:



Network	Nodes	Edges	N_{real}	λ
Gene regulation (transcription)				
<i>E. coli</i>	424	519	40	
<i>S. cerevisiae</i> *	685	1,052	70	
Neurons				
<i>C. elegans</i> †	252	509	125	
Electronic circuits (digital fractional multipliers)				

Concepte de bază

- ❑ Coeficient de clusterizare $C \rightarrow$ *are my friends, friends of my friends?*
 - Ia în calcul numărul de triunghiuri din rețea
 - Coeficientul unui nod = raportul dintre numărul legăturilor care conectează vecinii nodului și numărul de legături posibile între acești vecini
 - C - media coeficienților tuturor nodurilor din rețea



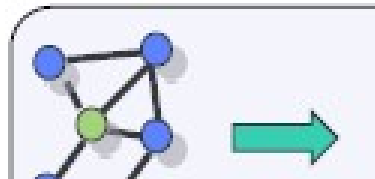
Concepte de bază

□ Eficiență globală

- Media armonică a drumurilor optime între toate nodurile rețelei

□ Eficiență locală (a unui nod) → coeficientul de clusterizare

- Lungimea celui mai scurt drum între mulțimea vecinilor nodului (fără a considera nodul)



Concepte de bază

□ Spectrul unui graf

■ Mulțimea valorilor proprii a matricii de adiacență sau a celei Laplaciene

□ Un graf cu N noduri are N valori proprii $\mu = (\mu_1, \mu_2, \dots, \mu_N)$ și N vectori proprii $V = (V_1, V_2, \dots, V_N)$.

□ importante pentru attributele topologice

- Diametru
- Număr de cicluri
- Propagarea informației

□ importante pentru attributele de conectivitate

■ Densitatea spectrală

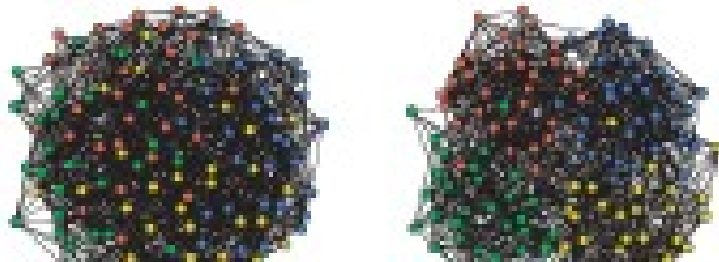
$$\rho(\mu) = \frac{1}{N} \sum_{i=1}^N \delta(\mu - \mu_i)$$

Concepte de bază

□ Structura comunităților

■ Comunitate în graf

- Un sub-graf ale cărui noduri sunt puternic conectate (sau cel puțin mai puternic conectate decât în cazul unei rețele aleatoare)

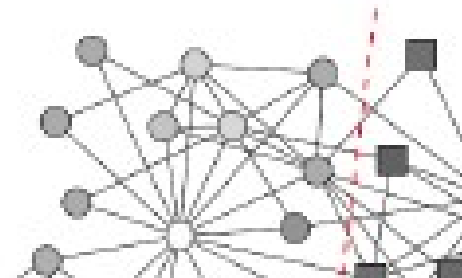


□ Evaluare

- Modularitate M

$$M = \sum_{i=1}^{N_M} \left[\frac{l_i}{l} - \left(\frac{d_i}{2l} \right)^2 \right]$$

Zachary Karate (



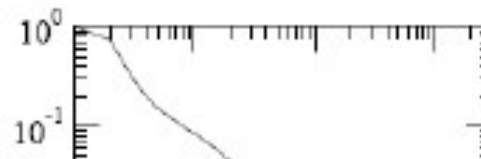
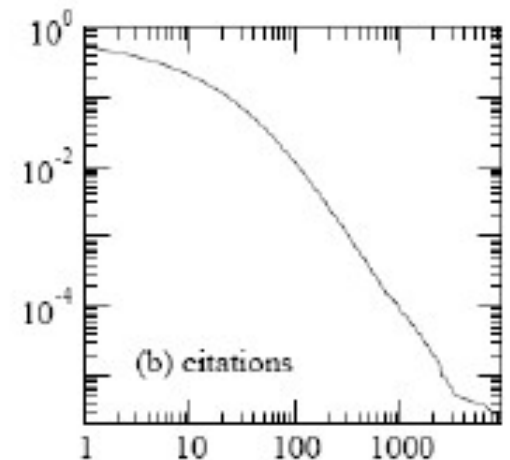
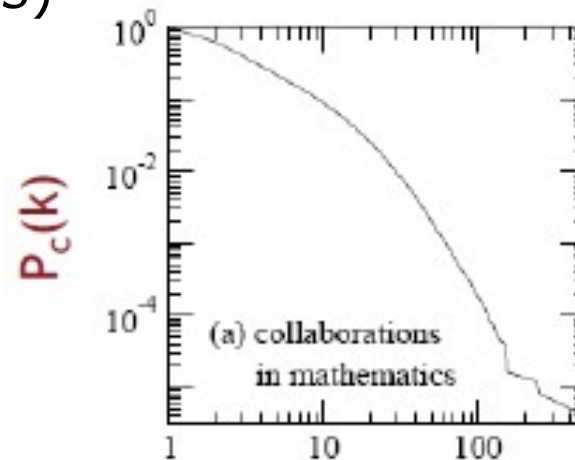
where N_M is the number of modules in the network, l_i is the number of links in module i , l is the total number of links in the network, and d_i is the sum of the degrees of the nodes in module i .

Concepte de bază

□ Distribuția gradelor

■ Distribuția (cumulată a) gradelor

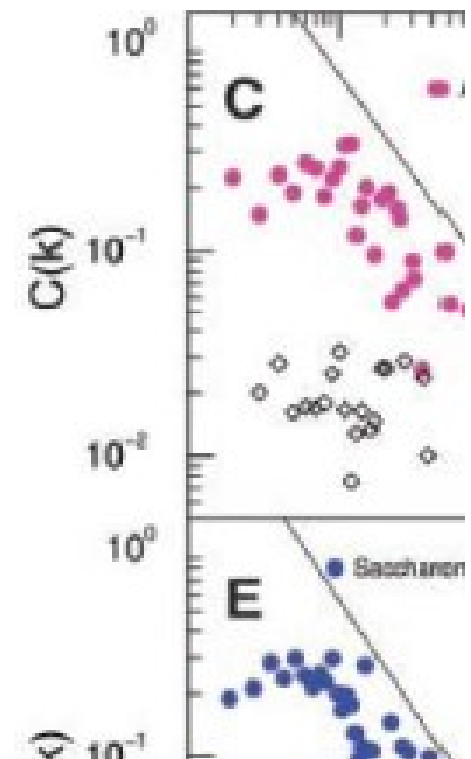
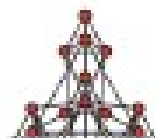
- Proporția nodurilor în rețea cu grad (mai mare sau) egal cu un prag k
- Poate fi exponențială (*random nets*) sau lege putere (*scale-free nets*)



Concepte de bază

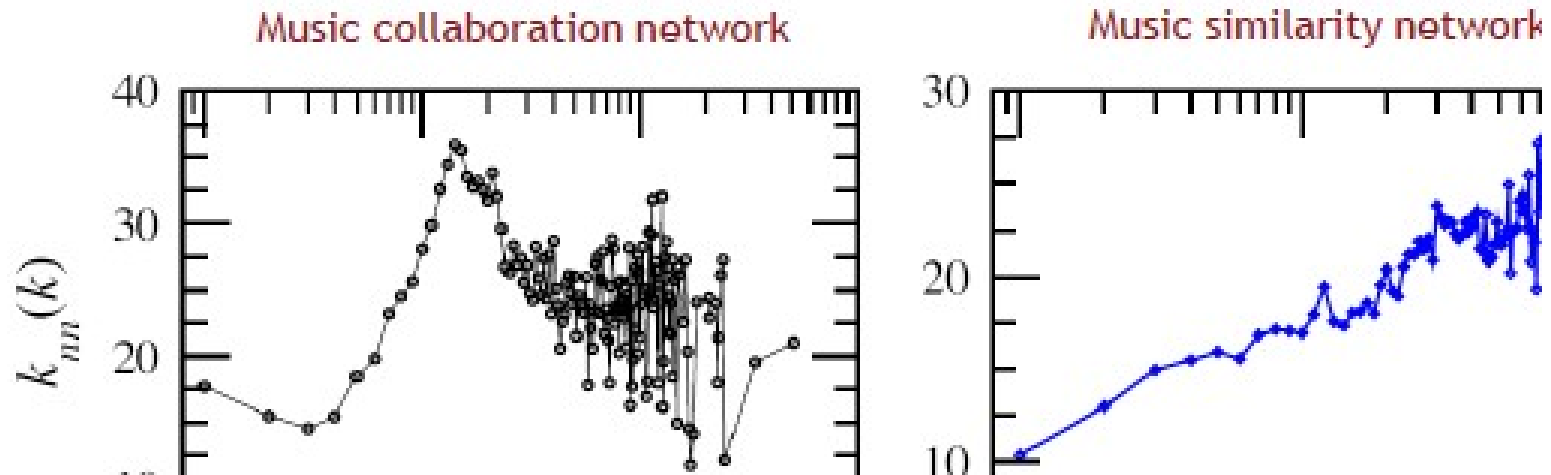
□ *Clustering distribution*

Figure: Clustering distribution in three organisms: *Aquidex aeolicus* (archaea) (C), *Escherichia coli* (bacterium) (D), and *Saccharomyces cerevisiae* (eukaryote) (E). (F) The $C(k)$ curves averaged over all 43 organisms is shown, and the inset displays all 43 species together. Lines correspond to $C(k) \sim k^{-1}$, and diamonds represent the $C(k)$ value expected for an equivalent scale-free network, indicating the absence of scaling



Concepte de bază

- *Nearest neighbor degree* $k_{nn}(k)$ și *assortativity*
 - $k_{nn}(k)$ – măsoara gradul vecinilor, fiind un indicator al asortării rețelei



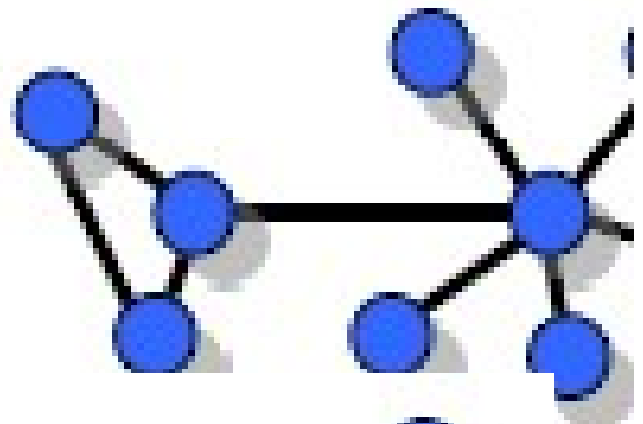
Tipuri de rețele complexe

- Direcția legăturilor
 - Neorientate
 - Orientate
- Tipul legăturilor
 - Neponderate
 - Ponderate
- Diferențele între noduri
 - Simple
 - Bipartite
- Topologie
 - Statice
 - Evolutive
- Dinamica nodurilor
 - Fără dinamică
 - Cu dinamică

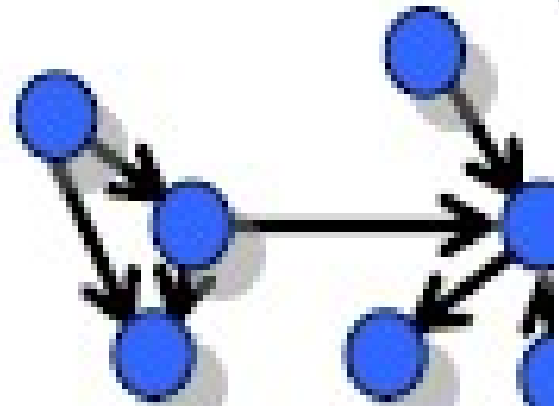
Tipuri de rețele complexe

□ **Direcția legăturilor** → procesele dinamice din rețea (propagarea informației, sincronizarea, robustețea)

■ Neorientate (legături simetrice)



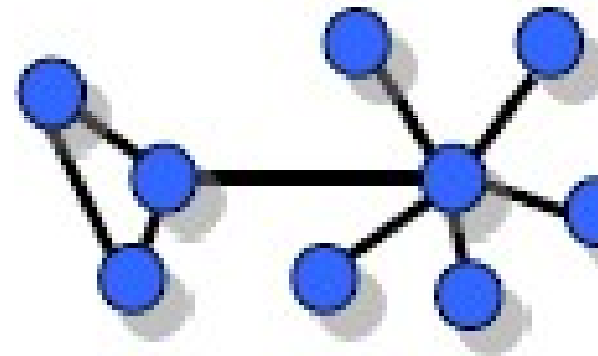
■ Orientate (legături asimetrice)



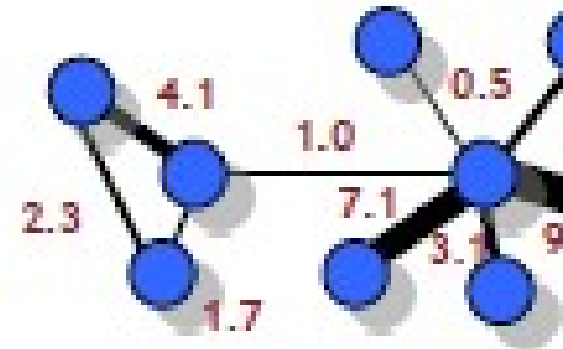
Tipuri de rețele complexe

□ **Tipul legăturilor** → procesele dinamice din rețea (propagarea informației, sincronizare, robustețe)

■ Neponderate (omogene)



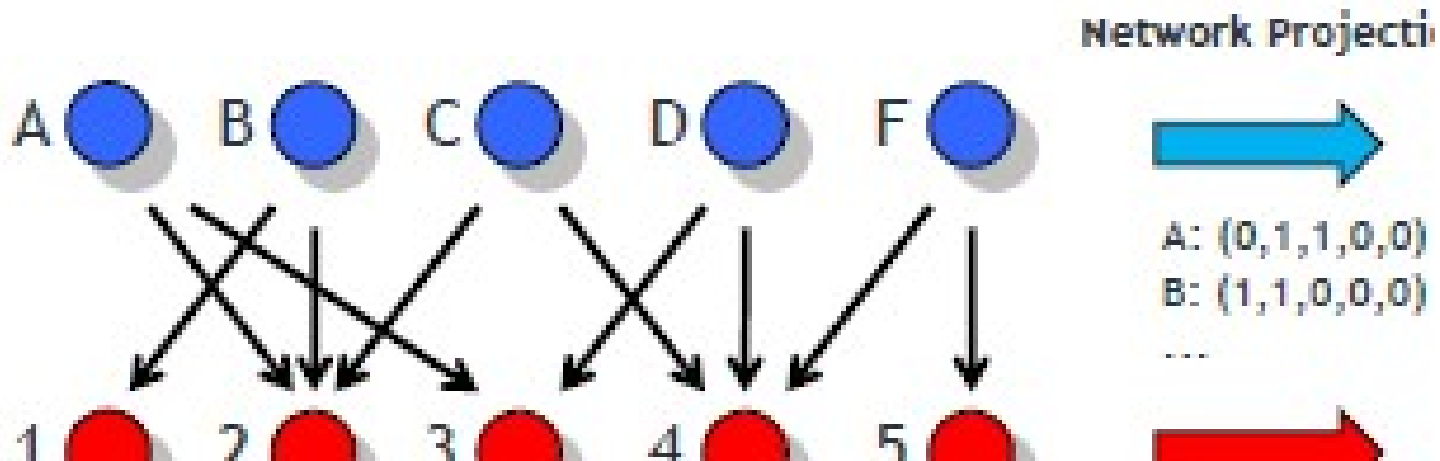
■ Ponderate (eterogene)



Tipuri de rețele complexe

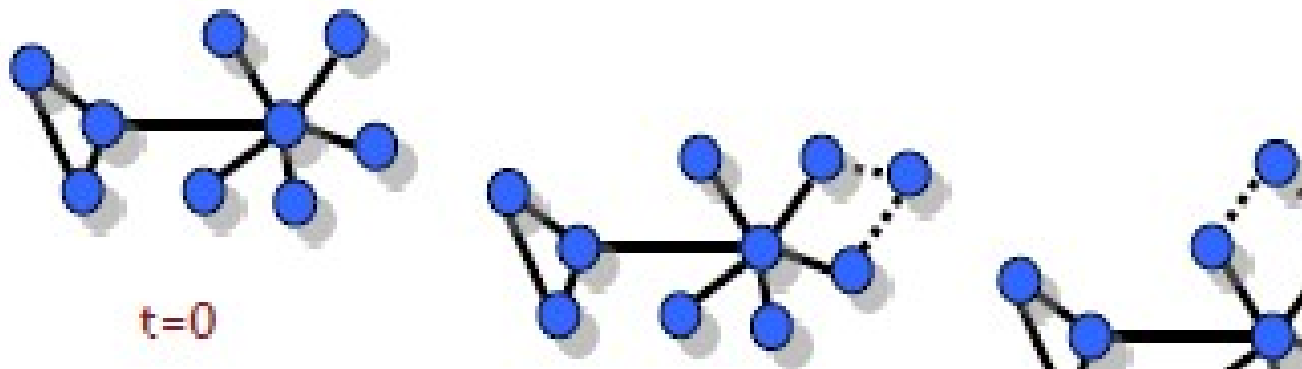
□ Diferențele între noduri

- Simple
- Bipartite → rețele cu noduri de 2 sau mai multe tipuri și legături **doar** între nodurile de același



Tipuri de rețele complexe

- **Topologie** → rețelele nu "apar" instantaneu;
 - Statice (ca și structură)
 - Evolutive (ca și structură)
 - Posibile întrebări:
 - Care sunt reguli ce guvernează evoluția?
 - Care sunt consecințele aplicării acestor reguli asupra topologiei finale a rețelei?



Tipuri de rețele complexe

□ Dinamica nodurilor

- poate fi influențată prin matricea conexiunilor;
- se poate analiza influența topologiei rețelei în procesele dinamice care apar în rețea
 - sincronizări,
 - procese stocastice
- dar și vice-versa (influența proceselor asupra topologiei rețelei)
- Fără dinamică
- Cu dinamică → nodurile sunt sisteme dinamice (cuplate):
 - oscilatoare periodice – ex. Pendulul lui Foucault,
 - sisteme excitabile,
 - oscilatoare haotice – ex. criptări,
 - sisteme bistabile – ex. *semiconductor memory*

Aplicații

□ Rețele sociale

■ Proprietăți

- Small-world
- Distribuții de tip power-low → noduri foarte conectate (*hub-uri*)
- Coeficient de clusterizare mare (relativ la rețelele aleatoare)
- Amestec de asortare
 - Ex: most connected nodes like to be together

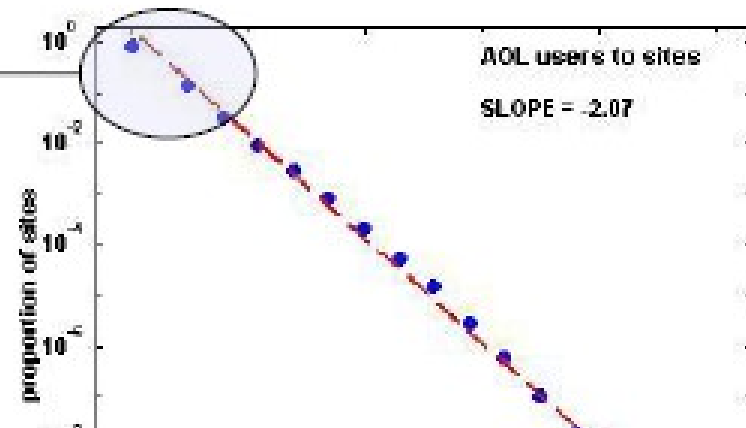
Aplicații

□ Rețele sociale

■ Vilfredo Pareto (1848 – 1923)

- Distribuția bunăstării (venituri, pământuri, etc.) în diferite țări
 - Au aceeași distribuție *power-law* $p(X \geq x) \sim x^{-\beta}$
- Principiu Pareto (regula 80-20):
 - 80% din efecte provin din 20% din cauze
 - Consecință a distribuției cumulate de tip *power law*
 - Aplicat în diferite domenii (economie ... sociologie)

Most sites have a low number of visits



Aplicații

□ Rețele sociale

■ George K. Zipf (1902 – 1950)

- Frecvența apariției cuvintelor în limba engleză urmează o distribuție a rangurilor (r) de tip *power-law*

- $N(r) \sim r^{-\gamma}$

- Contexte (la diferite scări)

- Cuvinte

- Silabe

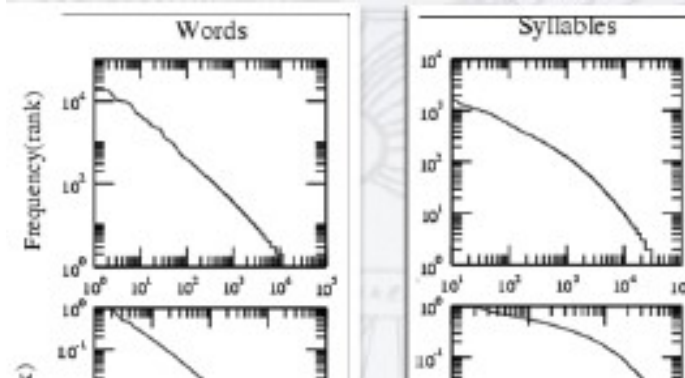
- Foneme

- Distribuția se schimbă atunci când se pierde semantica
 - Frecvența și gradul devin similare

Simon model: "as words are successively added to the context is created. As the context emerges, it favors the appearance of certain words in particular, those that have appeared – and inhibits the others."

$$n(r) = \frac{1}{(a + b)}$$

Redes Compleias → Redes



Aplicații

□ Rețele sociale

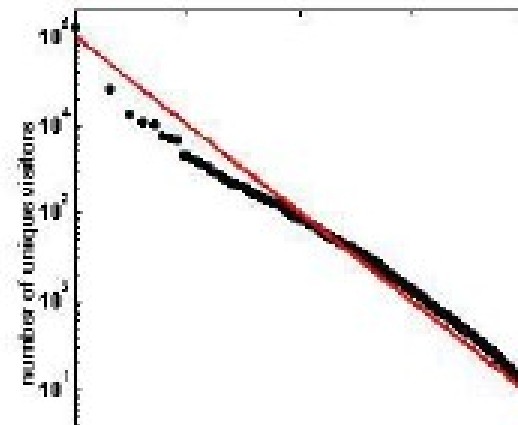
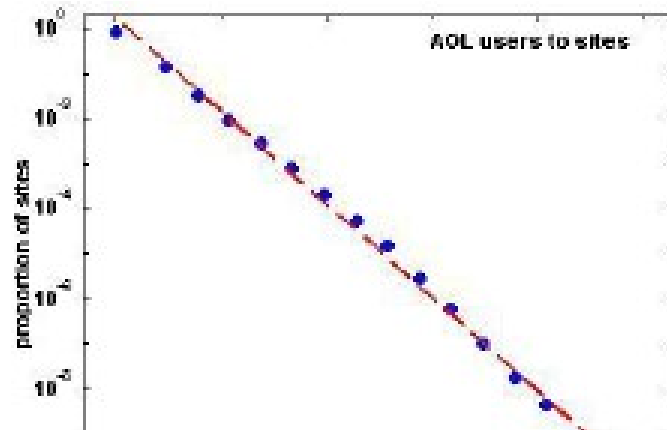
■ George K. Zipf (1902 – 1950)

□ Alte domenii

- Lingvistică
- Distribuția populației
- Rangurile veniturilor

□ Legătura cu Pareto

- al r -lea cel mai larg oraș are n locuitori (Zipf)
- r orașe au n sau mai mult de n locuitori (Pareto)



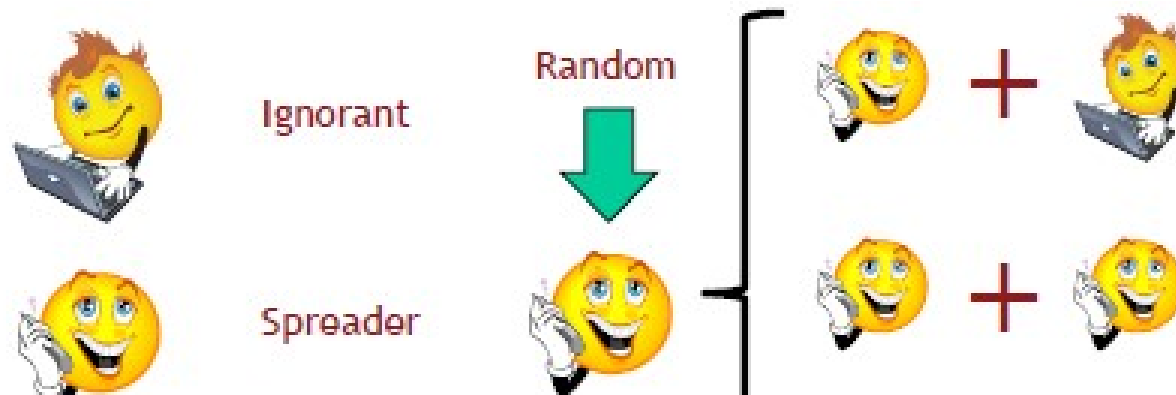
Aplicații

□ Rețele sociale

■ Răspândirea zvonurilor/bolilor

□ Modelul *Ignorant-Spreader-Stifler* (modelul Daley and Kendal)

- La fiecare moment de timp un informator aleator i este selectat; acesta contactează unul din vecinii săi j
 - Dacă j este ignorant, j se transformă în informator
 - Dacă j este un informator sau un ne-interesat, i se transformă în ne-interesat



Aplicații

□ Rețele sociale

■ Răspândirea zvonurilor/bolilor într-o rețea small-world

- Modelul Ignorant-Spreader-Stifler (modelul Daley and Kendal)

After a certain transient, the rumor stops:

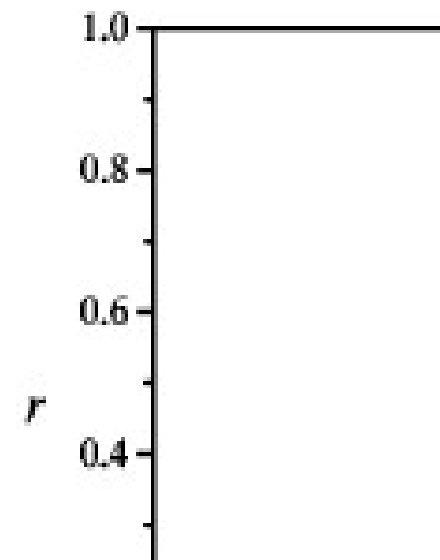
N = total number of nodes

$r = N_{\text{stiflers}}/N$

P = probability of rewiring

$\lambda = \beta = \alpha$ (for the case of the figure)

- Below a certain p_c the rumor does not



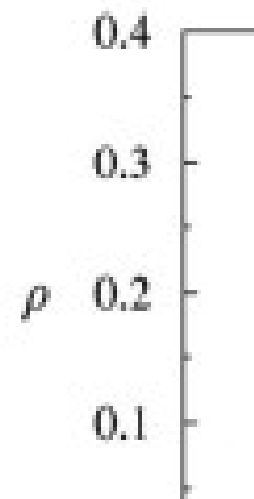
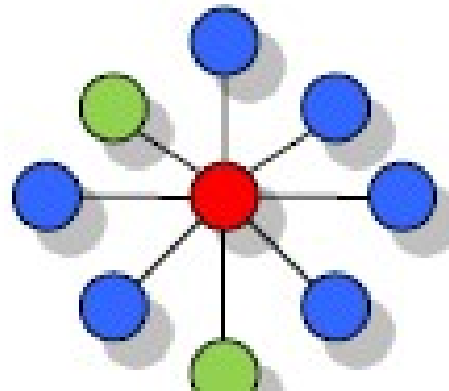
Aplicații

□ Rețele sociale

- Răspândirea zvonurilor/bolilor într-o rețea small-world
 - Cum să luptăm împotriva răspândirii?

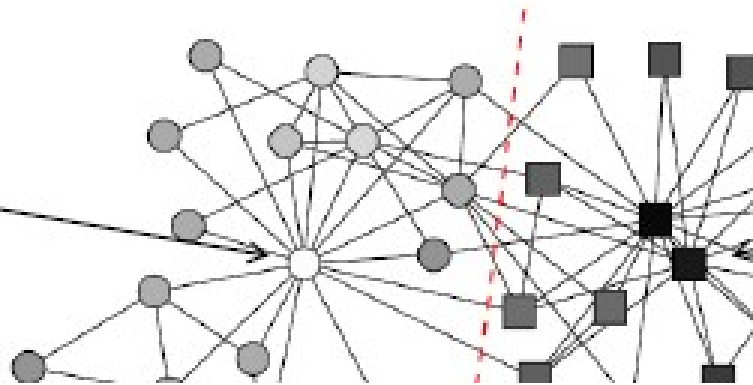
λ = spreading rate

g = fraction of immune nodes



Aplicații

- Rețele sociale
 - Comunitățile și rolul lor
 - Clubul Zachary (de karate)



Aplicații

- Rețele sociale → Comunitățile și rolul lor
 - Structura modulară a rețelelor complexe
 - Existența comunităților nu se reflectă în distribuția gradelor, clustering sau assortativity
 - Comunitățile sunt relaționate de funcțiile nodurilor
 - Detectia comunităților = problemă cu multiple soluții
 - Algoritmi
 - Metode de eliminare a legăturilor
 - Metode aglomerative
 - Algoritmi bazați pe modularitate
 - Metode spectrale

Aplicații

- Rețele sociale → Comunitățile și rolul lor → Detectia comunităților
 - Metode de eliminare a legăturilor
 - Se tot elimină legăturile slabe până când rețeaua începe să se rupă

Example:

The weakest links are those with the highest betweenness.



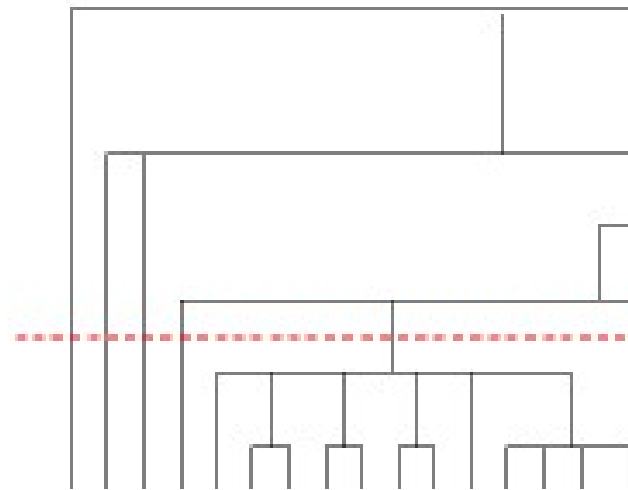
Aplicații

- Rețele sociale → Comunitățile și rolul lor → Detecția comunităților
 - Metode aglomerative (*Bottom-up*)
 - Inițial fiecare nod aparține unei comunități
 - Nodurile similare

Example:

The similarity is based on the structural equivalence

$$x_{ij} = \sqrt{\sum_{k \neq i,j} (A_{ik} - A_{jk})^2}$$

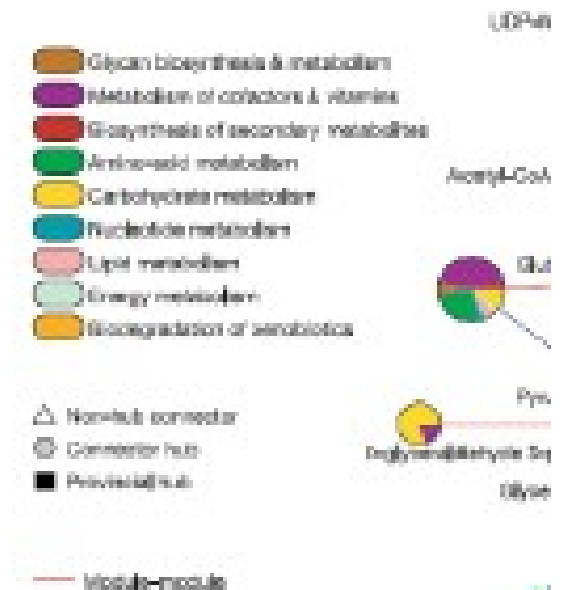


Aplicații

- Rețele sociale → Comunitățile și rolul lor → Detecția comunităților
 - Algoritmi bazați pe modularitate
 - Modularitatea = numărul de elgături dintre grupuri – numărul estimat de legături între grupuri într-o rețea aleatoare

Example:

The simulated annealing algorithm is an stochastic algorithm that searches for the maximum modularity by allowing negative (decreasing) variations of the modularity.



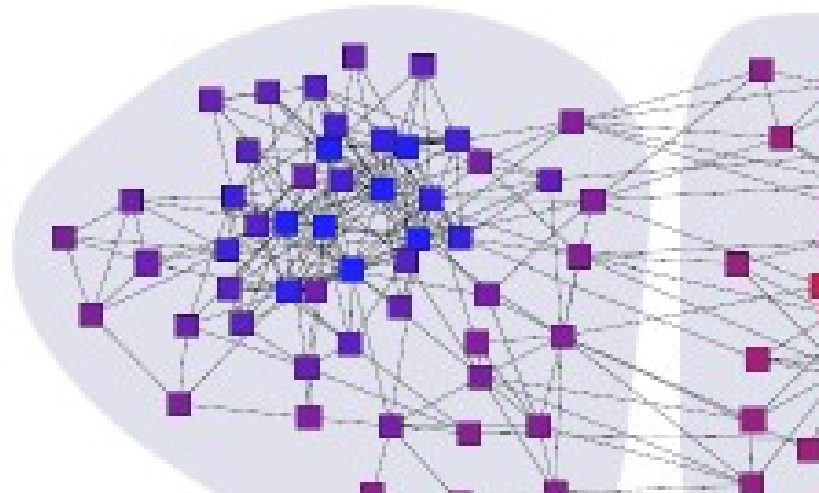
Aplicații

□ Rețele sociale → Comunitățile și rolul lor → Detectia comunităților

■ Metode spectrale

□ Separarea rețelei în

- 2 (sau mai multe) componente
- prin analiza spectrală a matricii Laplaciene
 - care conține toate informațiile topologice despre rețea



Aplicații

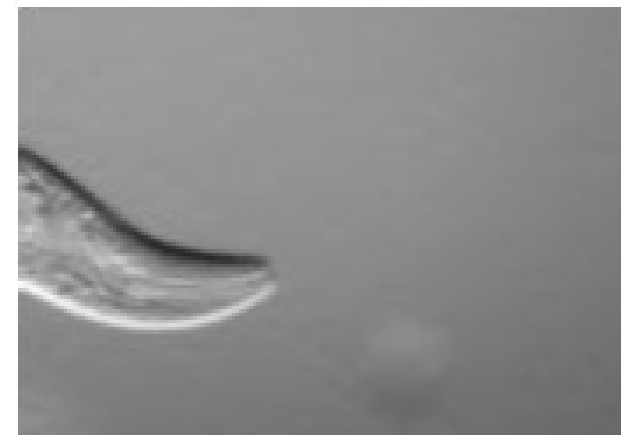
□ Rețele biologice

■ Proprietăți

- Small-world
- Dissortative mixing
 - multe dintre nodurile conectate nu sunt conectate preferențial unele cu altele
- Organizate în sub-module
 - Modularitate ridicată
 - Structuri comunitare

■ Pionieri

- Watts și Strogatz



L_{actual}

L_{random}

File: ...

0.05

0.05

Aplicații

□ Rețele biologice

■ Rețele metabolice

■ Rețele proteice

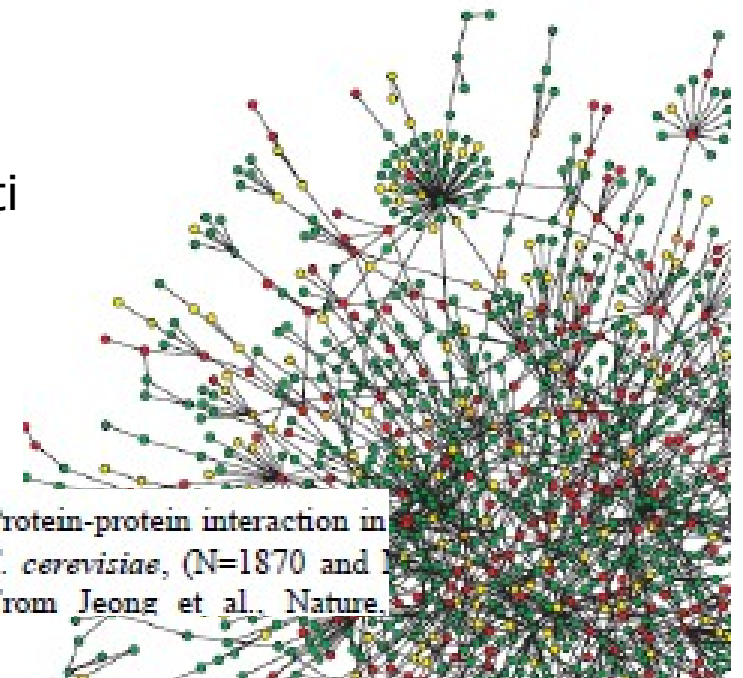
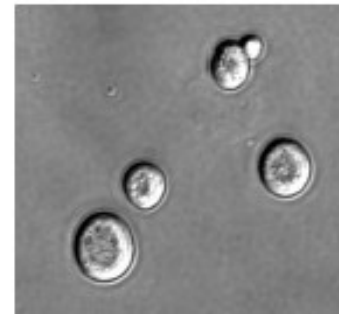
□ Drojdia

- 6000 proteine
- 3 interacțiuni per proteină
- 20 000 interacțiuni

□ Corpul uman

- De ordinul 100 000 de interacțiuni

■ Rețele genetice



Protein-protein interaction in *S. cerevisiae*, (N=1870 and E=1870)
From Jeong et al., Nature.

- Rețele neutre ale acidului ribonucleic (RNA)

- Aceeași structură poate fi obținută cu o multitudine de diferite lanțuri

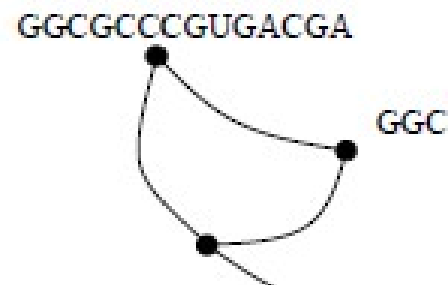
C: cito

$$(((\dots(((((\dots))))))\dots))\dots$$

AGCAAGUGCAGUUGCACAAGGAUCUCAUCCAGCU

(((---(((---)))---)))---(((---)))---

Figure 6

[illegible][illegible]

Aplicații

□ Rețele biologice

■ Rețele funcționale ale creierului

□ RMN funcțional

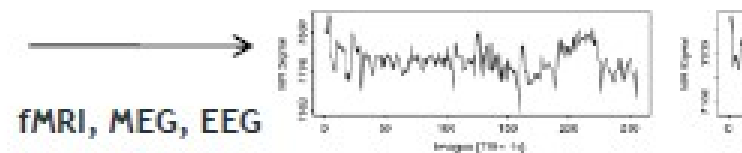
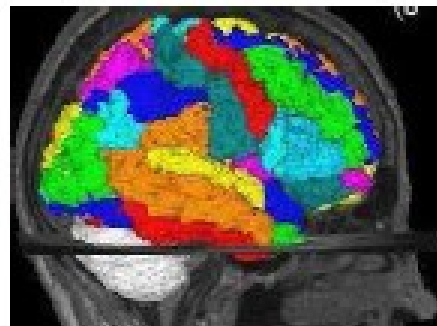
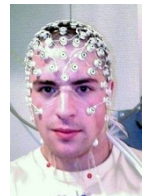
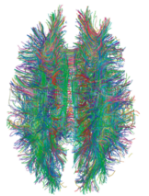
- Măsoară cum diferite părți ale creierului răspund la stimuli exteriori sau la activități pasive într-o stare de repaus

□ Electroencefalogramă

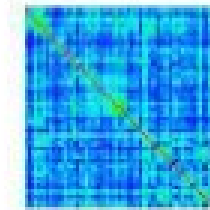
- înregistrarea biocurenților câmpurilor transcranieni la nivelul tegumentelor și înregistrarea grafică a acestora în timp sub forma de unde

□ Magnetoencefalogramă

- tehnică de imagistică neurală funcțională pentru cartografierea activității **creierului** prin înregistrarea câmpurilor magnetice produse de curenții electrici ce apar în mod natural în creier,



Correlation
Matrix



Grupo de Sistemas Evolutivos (C.A.B.)

Aplicații

□ Rețele biologice

■ Rețele funcționale ale creierului

- Alzheimer
- Mild Cognitive Impairment
- Schizofrenie
- Epilepsie

Aplicații

□ Rețele muzicale

- Rețele ale notelor
- Rețele ale melodiilor
- Rețele ale artiștilor
- Rețele ale utilizatorilor

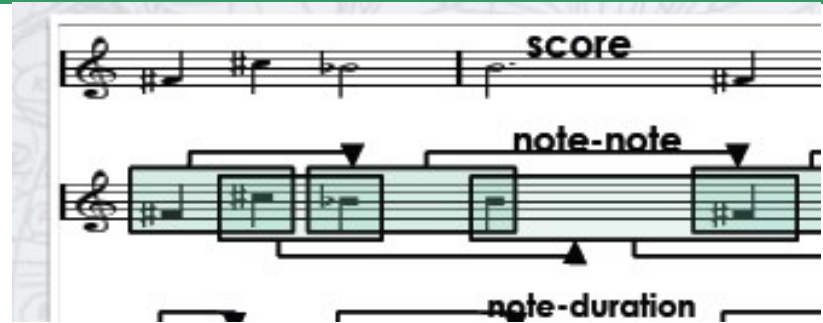


Aplicații

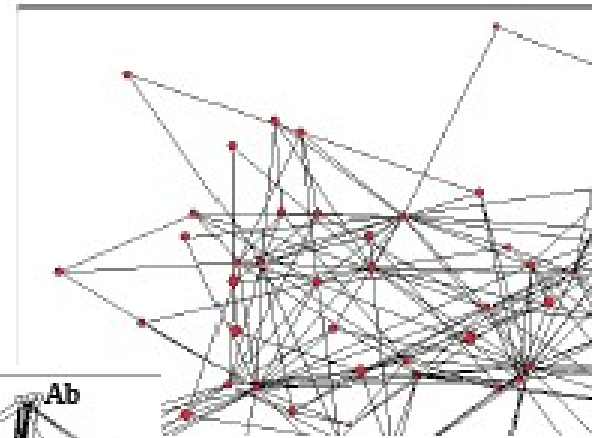
□ Rețele muzicale

■ Rețele ale notelor

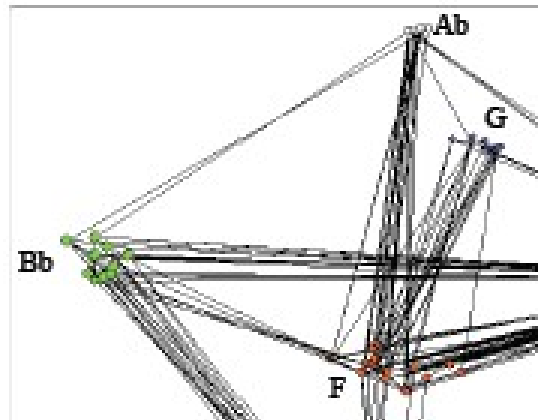
- Nota = nod
- Legătura = proximitatea (apropierea) notelor



□ Rețeaua duratei notelor



□ Rețeaua notelor



Aplicații

□ Rețele muzicale de note

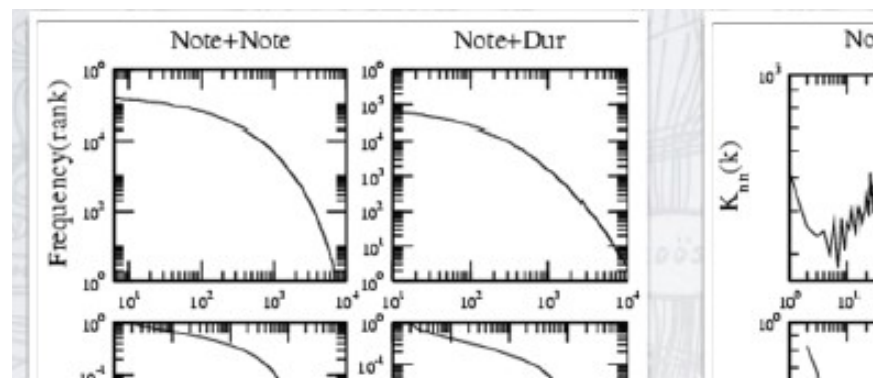
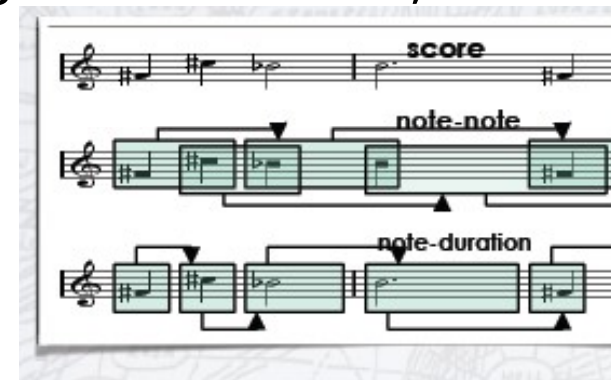
■ Legea Zipf

- Context = o ierarhie a modelelor (șabloanelor) care apar în diferite tempo-uri (progresie armonică, melodie, ton, ritm, ...)

- Note + durată
- Note + note

■ Rețelele de note

- sunt assortative
- au o distribuție de clusterizare specifică



Aplicații

□ Rețele muzicale de note

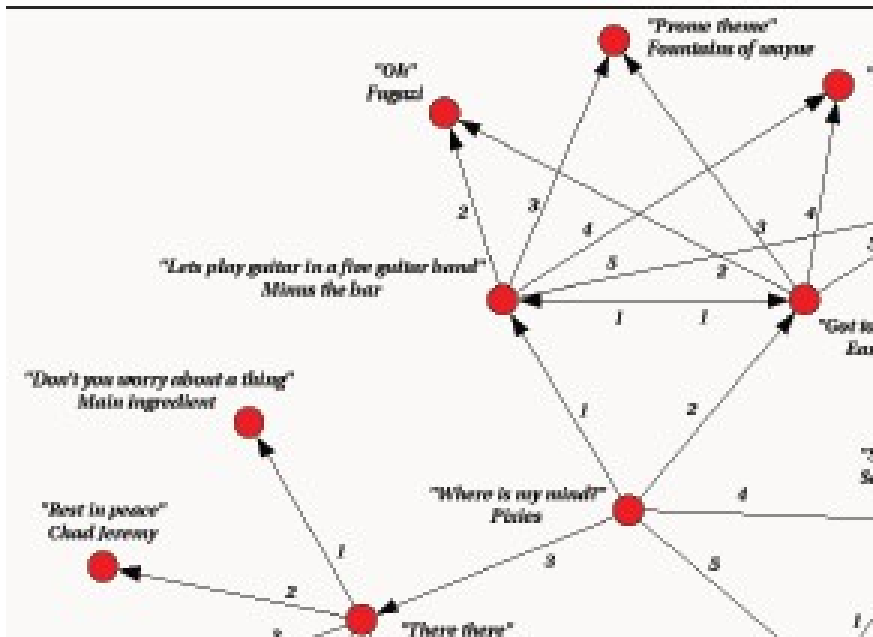
■ De ce?

- Sa creăm muzică pornind de la proprietățile rețelei și folosind procese aleatoare ghidate



- Rețele ale melodiilor

- Nod = melodie
- Legătura = diferite relații
 - Ex. Apariția în playlist-uri



From "The complex musical tastes", J.M. Buldú
Kronenhorn LA Almor

Aplicații

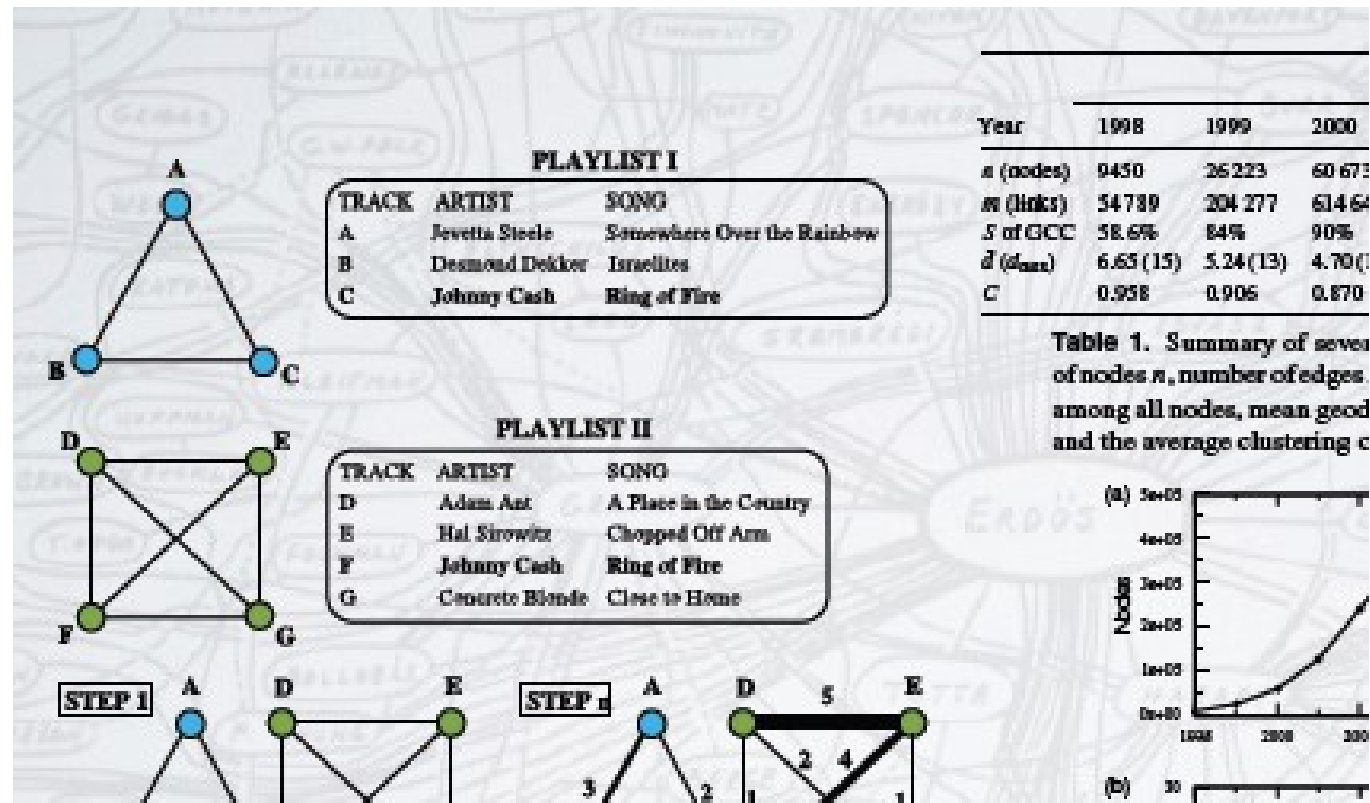
□ Rețele muzicale ale melodiilor

■ De ce?

- Analiza structurii
 - Drumuri, module
- Detectarea celei mai influente melodii
- Clasificare (etichetare)
- Proiectarea unor sisteme automate de recomandare eficiente

Aplicații

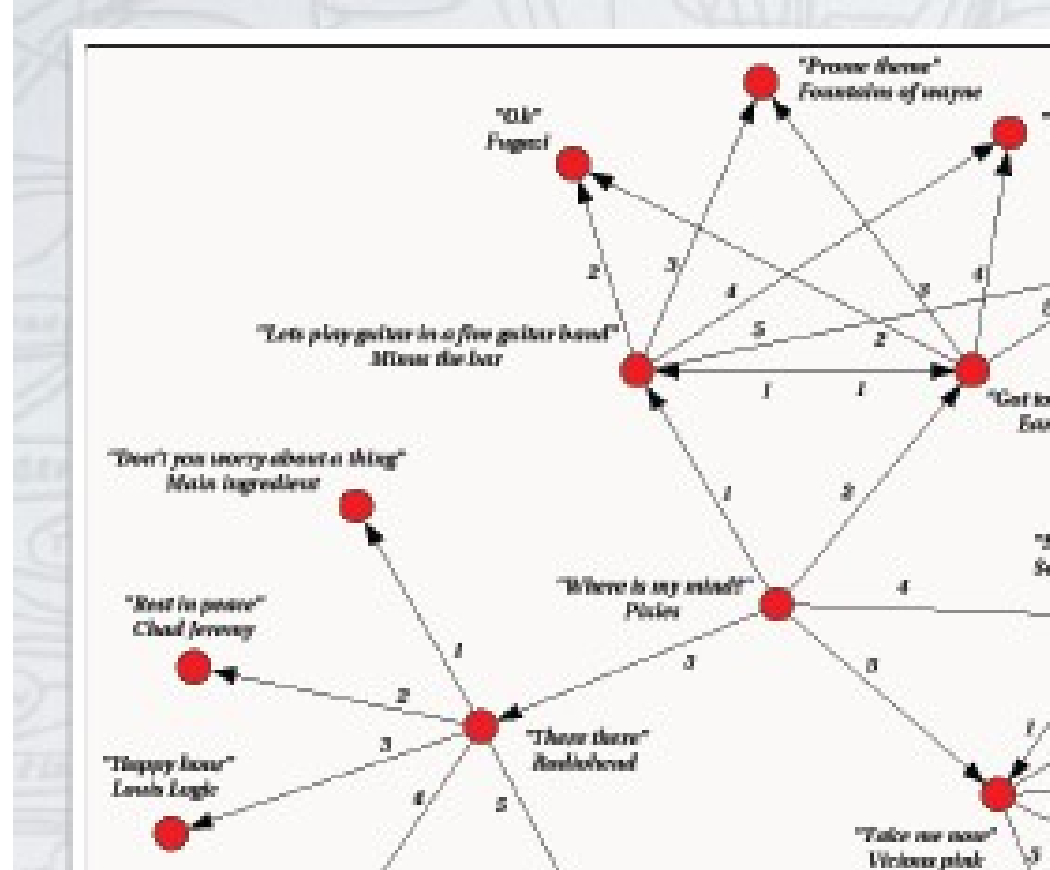
- Rețele muzicale de melodii
 - Evoluează în timp



□ Rețele muzicale de melodii

■ Gusturi musicali

- ## ■ Apartenența melodiilor la playlist-urile utilizatorilor

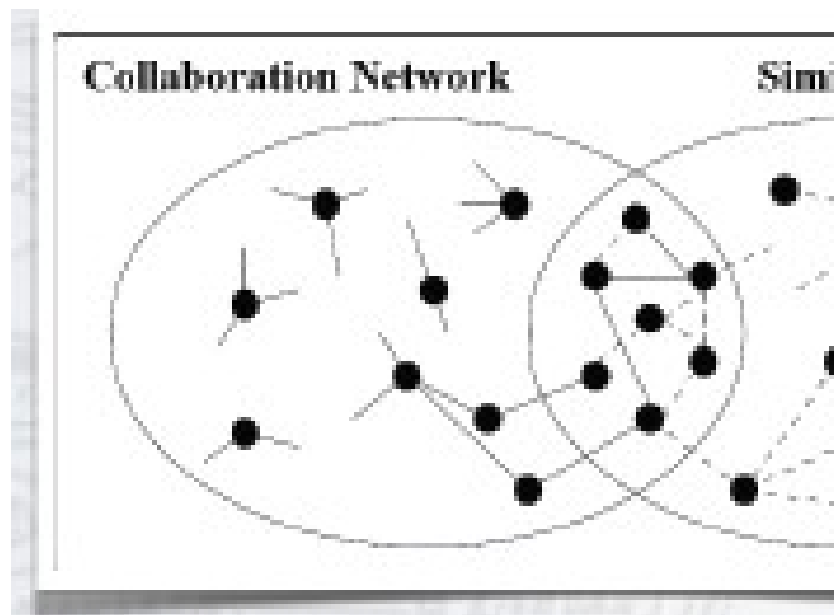
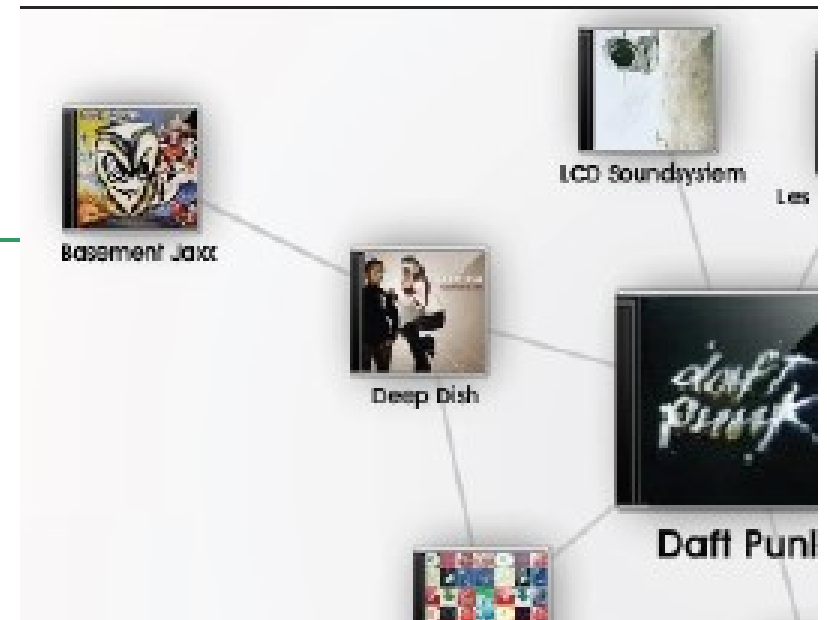


Aplicații

□ Rețele muzicale

■ Rețele ale artiștilor

- Nod = artist
- Legătură = diferite relații
 - Similaritate
 - Colaborare
 - Afinitate

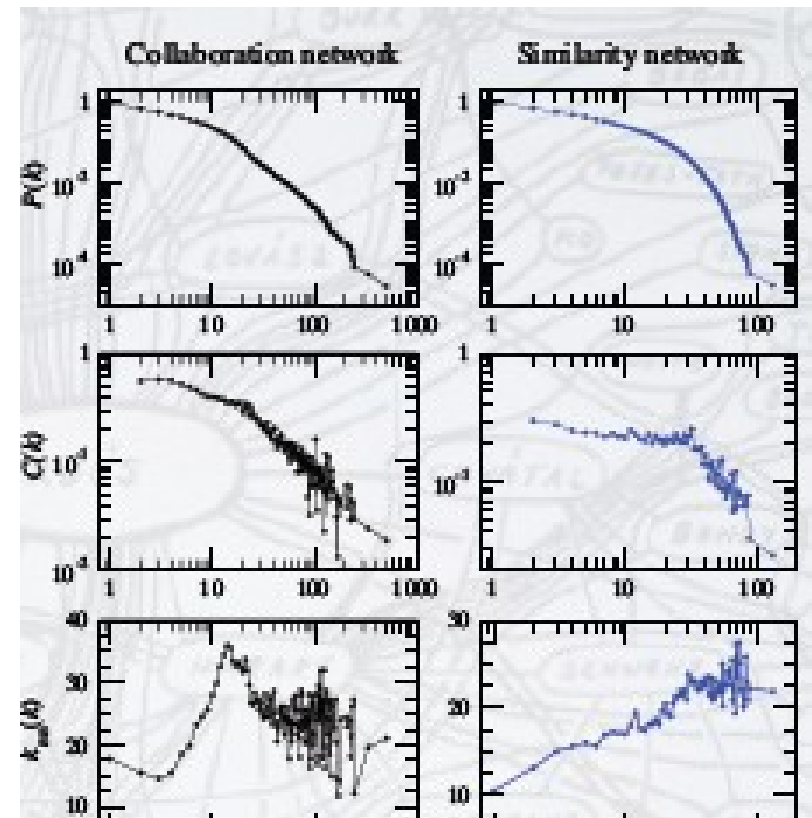


Aplicații

□ Rețele muzicale

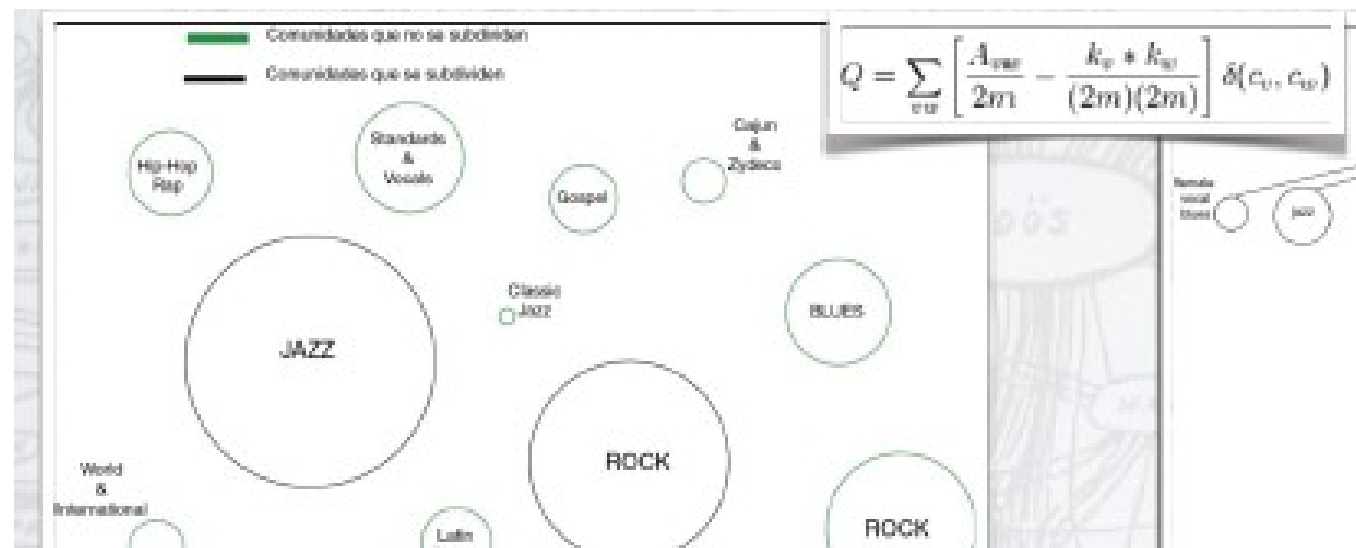
■ Rețele ale artiștilor

- Nod = artist
- Legătură = diferite relații
 - Similaritate
 - Colaborare
 - Afinitate
- Au același număr de noduri
- Au structuri diferite



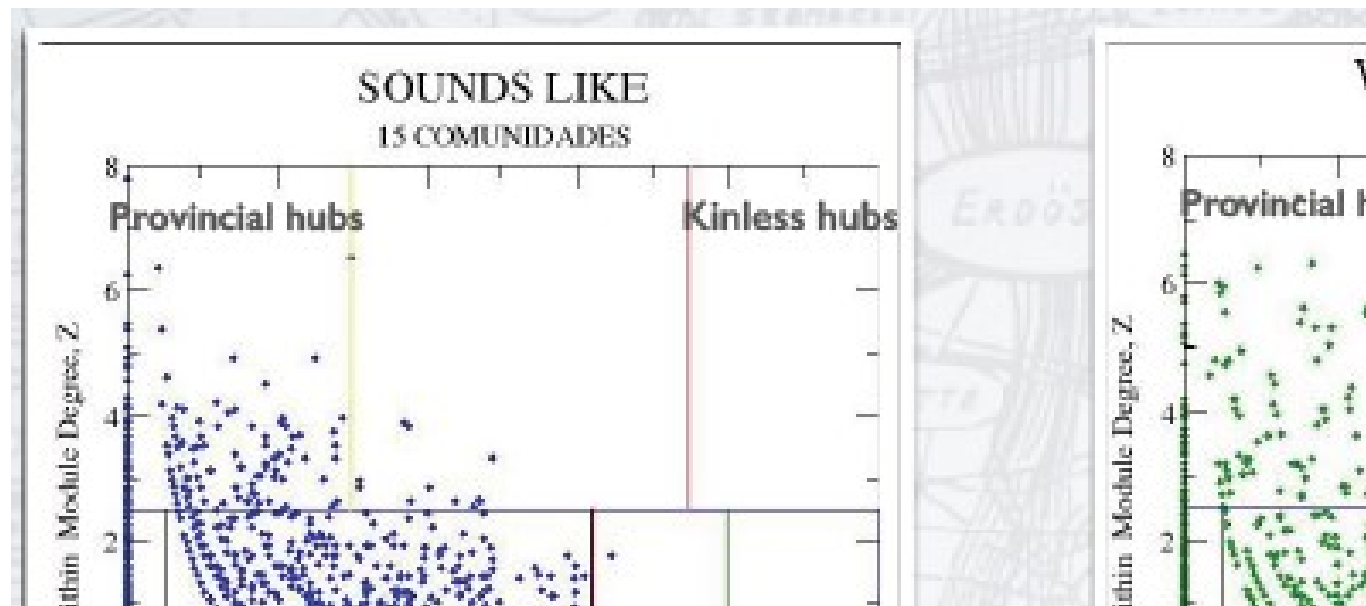
Aplicații

- Rețele muzicale ale artiștilor
 - Detecția comunităților oferă informații despre rețea
 - Separarea rețelei pe baza modularității
 - În comunități
 - În sub-comunități



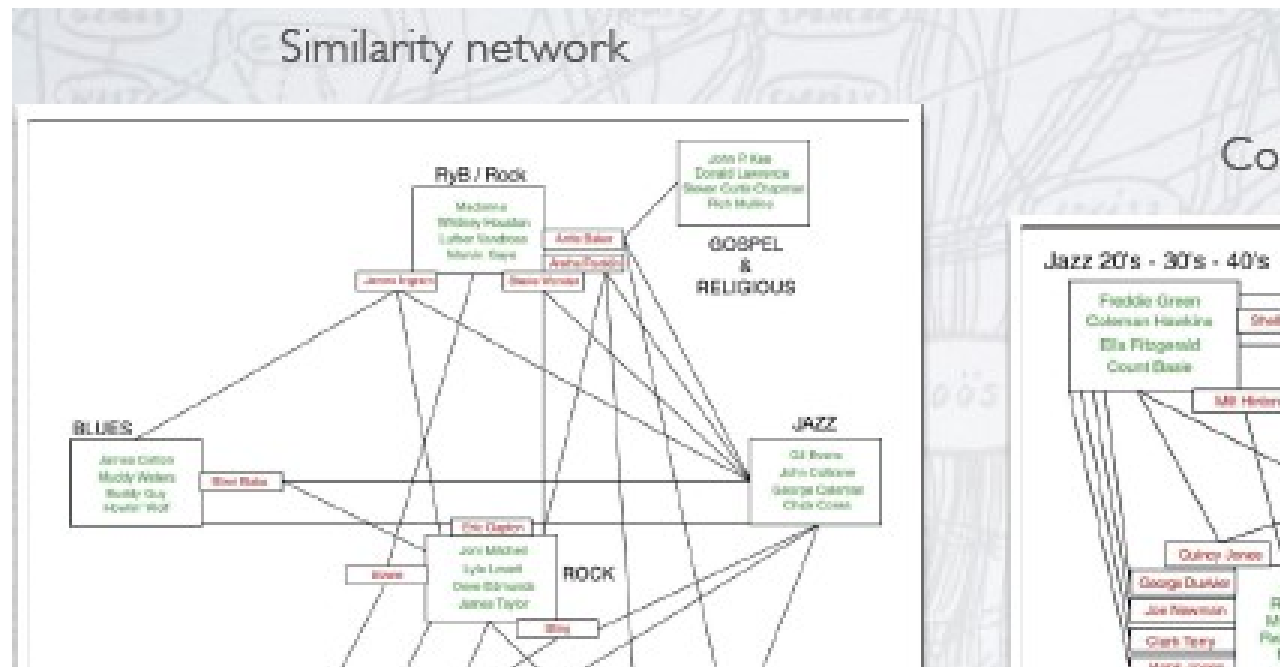
Aplicații

- Rețele muzicale ale artiștilor
 - Detecția comunităților oferă informații despre rețea
 - Separarea rețelei pe baza modularității
 - Se pot identifica hub-uri



Aplicații

- Rețele muzicale ale artiștilor
 - Detecția comunităților oferă informații despre rețea
 - Separarea rețelei pe baza modularității
 - Se pot identifica hub-uri
 - Se pot face cartografieri ale hub-urilor

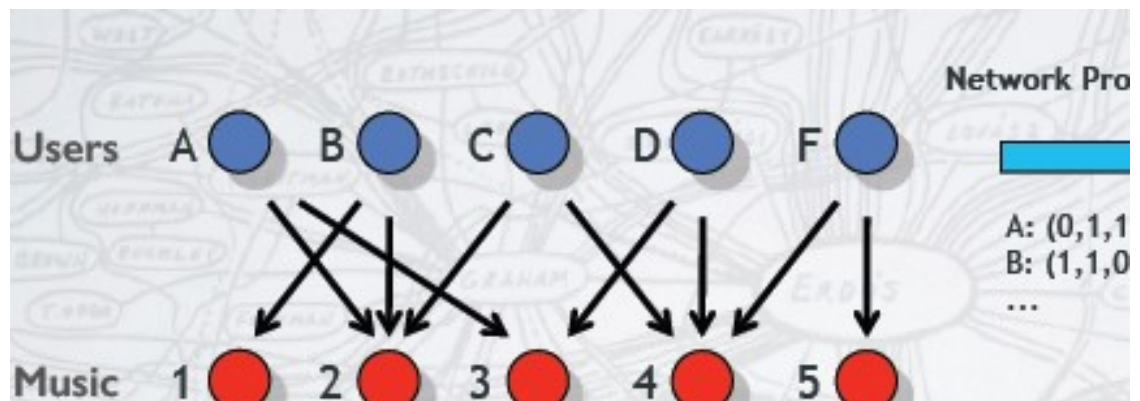


Aplicații

□ Rețele muzicale

■ Rețele ale utilizatorilor

- Nod = utilizator (consumator)
- Legătură = diferite relații



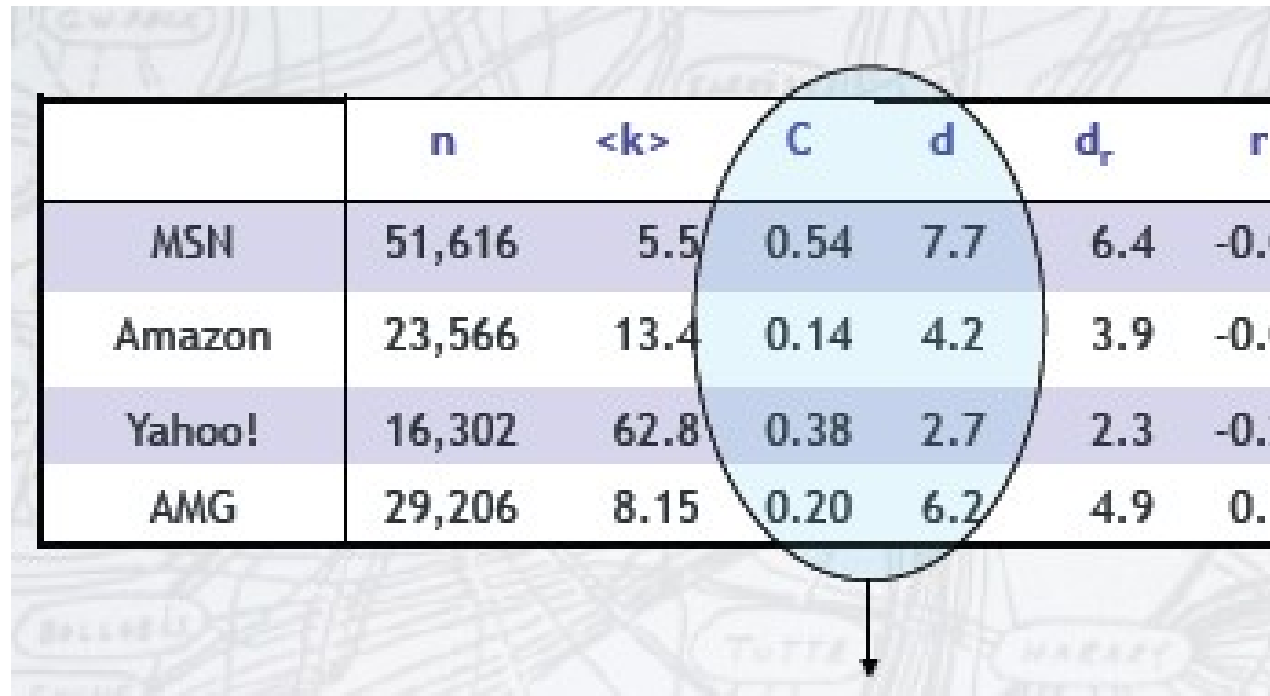
■ De ce?

- Folosite în sistemele de recomandare



Aplicații

- Rețele muzicale ale utilizatorilor
 - De ce e importantă topologia rețelei?



	n	<k>	C	d	d _r	r
MSN	51,616	5.5	0.54	7.7	6.4	-0.
Amazon	23,566	13.4	0.14	4.2	3.9	-0.
Yahoo!	16,302	62.8	0.38	2.7	2.3	-0.
AMG	29,206	8.15	0.20	6.2	4.9	0.

Aplicații

□ Rețele muzicale

- O diversitate de a proiecta muzica într-o rețea
- Care model de rețea este cel mai potrivit?

