

The background features a close-up of a laptop keyboard with a blue color scheme. Overlaid on the keyboard is a glowing globe of the Earth, surrounded by a network of white lines and nodes, symbolizing global connectivity and the Internet. The overall aesthetic is futuristic and digital.

# **Az Internet ökoszisztémája és evolúciója**

**Rétvári Gábor, Heszberger Zsolt**

# Tartalom

- Ismétlés korábbról – Az Internet mint szervezett káosz alapvető jellemzői
  - Nagyméretű, sokcsomópontos hálózat, komplex dinamika, skálázódási kérdések
- Nagy hálózatok elmélete – Bevezetés
  - Mivel foglalkozik a hálózattudomány?
- Valós hálózatok és hálózati modellek
  - Hálózati mérések
  - Hálózati modellek
- Hálózatdinamika: Internet, mint nagy hálózat
  - Hálózati alkalmazások
  - Folyamatok nagy hálózatokon

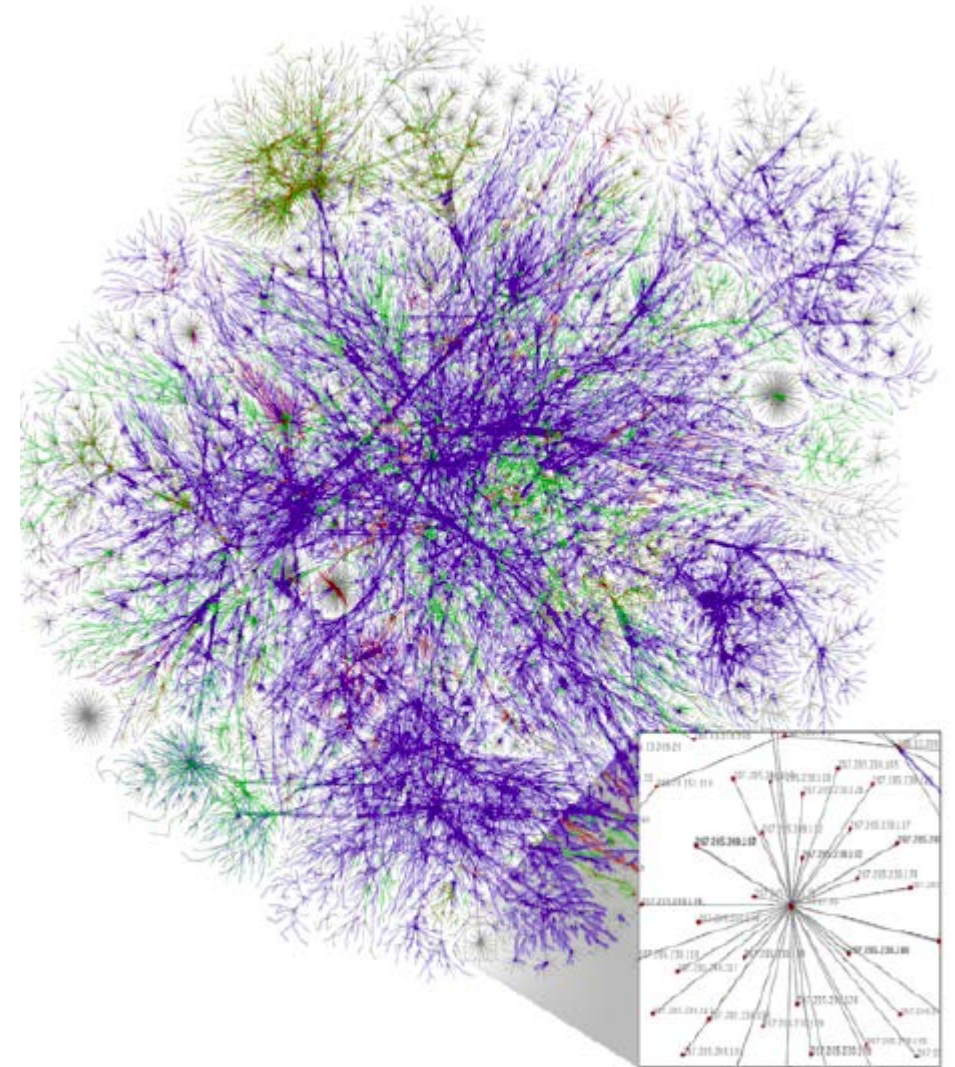


# Internet: A szervezett káosz



# Az Internet: „hálózatok hálózata”

- 50 ezer szolgáltató
- 10 milliárd csatlakoztatott eszköz
- 3 milliárd felhasználó
- Több 100 milliárd USD üzleti bevétel
- Kulcs kérdések:
  - Irányíthatóság? (hogyan és ki?)
  - Megbízhatóság?
  - Biztonság?
  - Magánszféra védelme?
- Mit hoz a jövő?

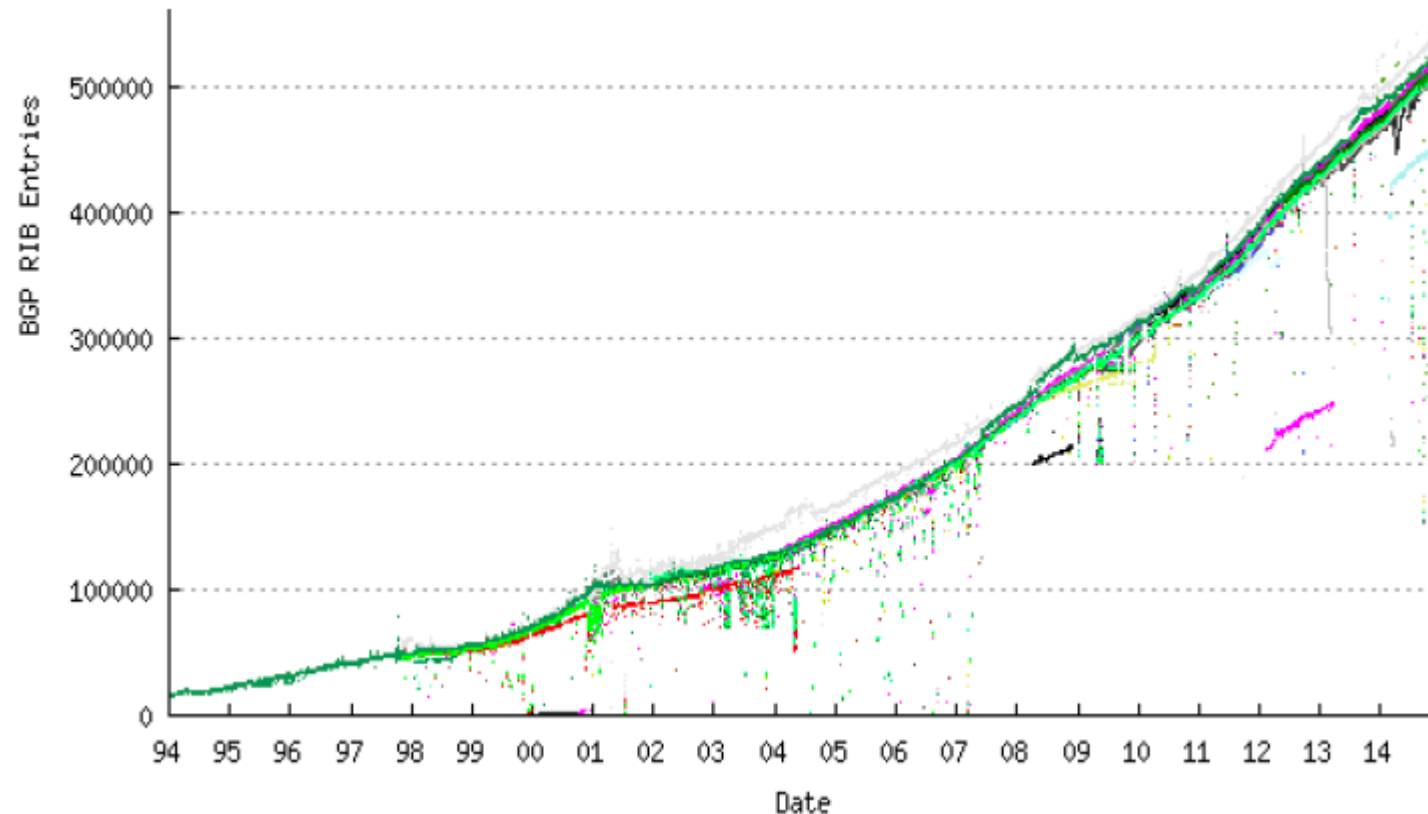


Opte Project, Wikipedia



# Skálázhatóság: A routing táblák növekedése

- Egyre több csatlakoztatott eszköz
- Egyre több IP címet kell tárolni a routing táblákban (az aggregáció ellenére)



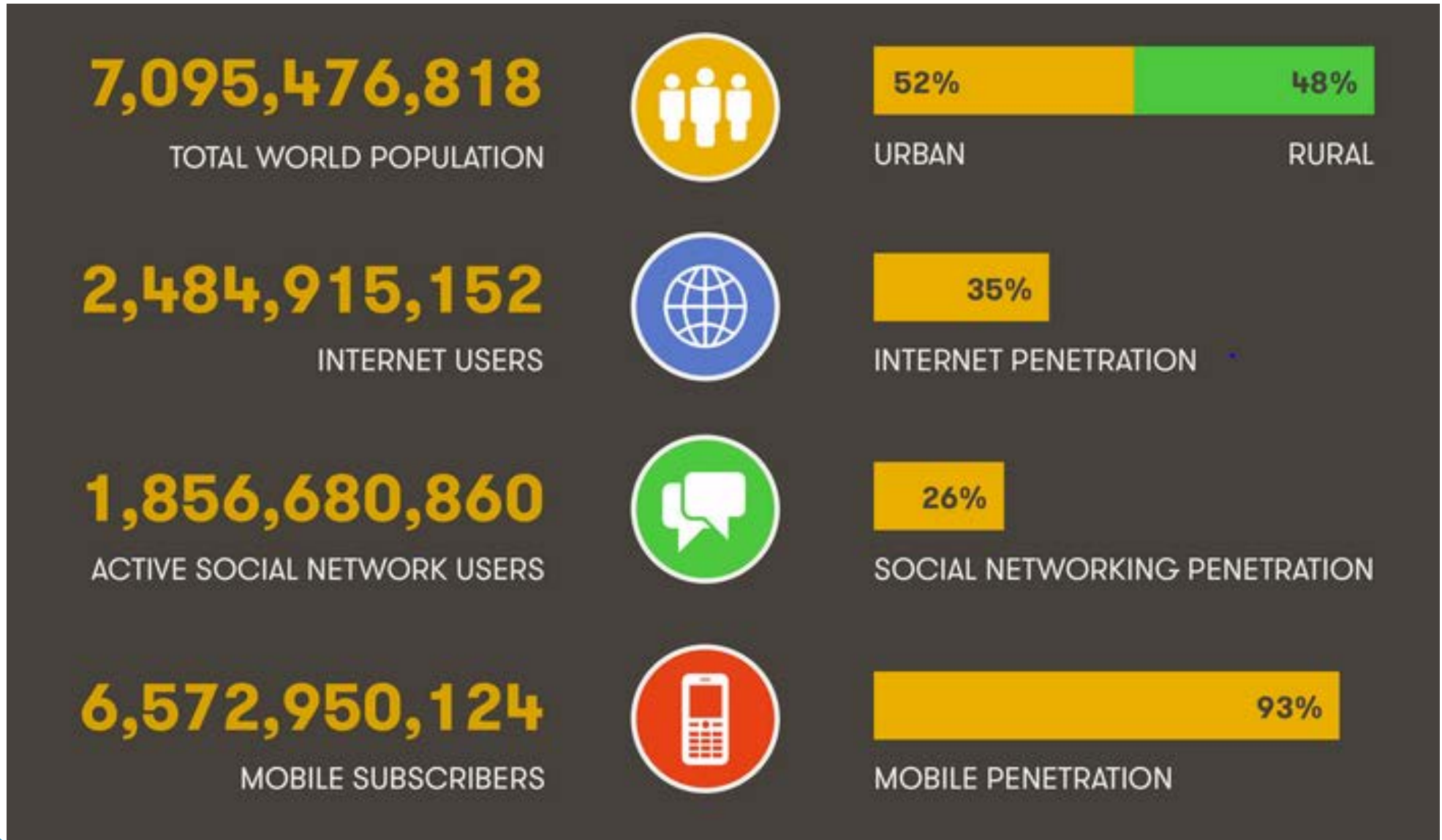
# Megoldások a skálázhatóságra?

- Konfiguráció
  - a tábla fizikai méretkorlátja sokszor igazából 1024k
  - de 512k bejegyzés foglalt az IPv6 számára
  - elég a fenti alapértelmezett korlátot átállítani
  - mi lesz, ha elérjük az 1024k bejegyzést?
  - és mi van a sebességgel?
  - meddig skálázhatóak a routerek hatékonyan?
- Új protokollok, új routing architektúra?

*Hogy lehet, hogy ilyen kritikus rendszer, mint az internet, már rövid távon sem skálázódik?*



# Internetes közösségi hálók aktív felhasználói? (2014.01)



# Kutatás: Komplex rendszerek vizsgálata: hálózatelmélet

- Komplex rendszer: Az építőkövek, közreműködők részletei lényegtelenek, lényeg a közöttük lévő kapcsolat
- Hálózatos látásmód: Az építőkövek egy általános gráf/hálózat részei, melyeket kapcsolatok/élek kötnek össze
- A rendszer segítségével összegyűjthető adatok korábban nem ismert skálán működő törvényszerűségeket világítanak meg



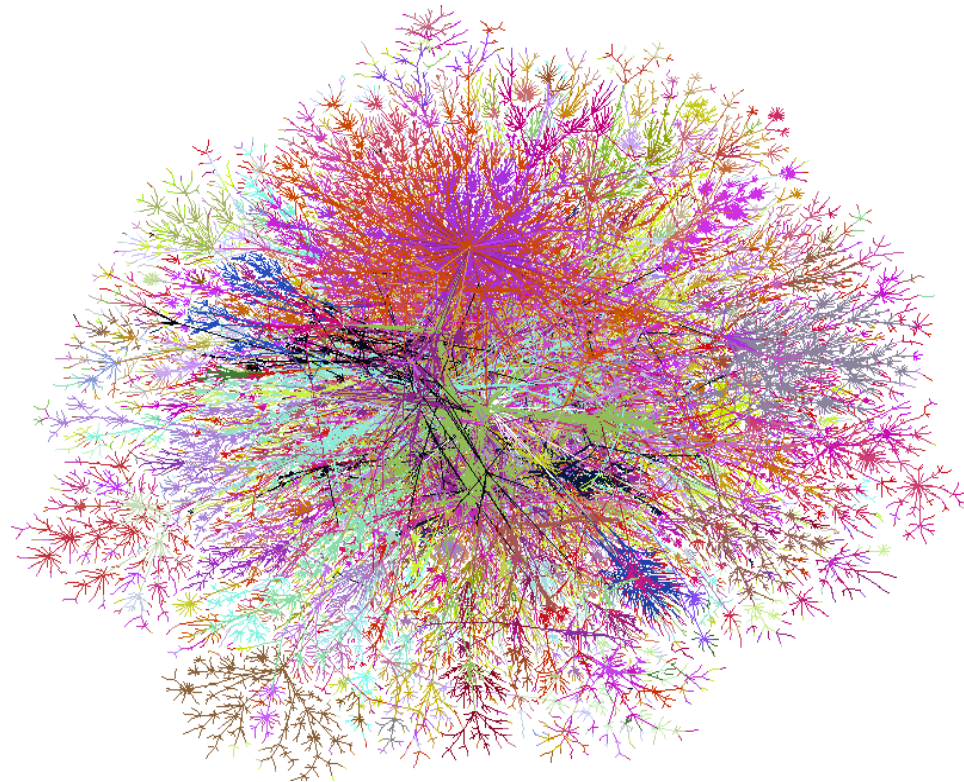


# A komplex hálózatok tudománya

- A komplex hálózatok tudománya
- Kialakulása 2000 körülre tehető, amikor feltűnt a kutatóknak, hogy nagy valós hálózatok nem teljesen véletlenek, mint azt korábban feltételezték.
- A kutatás fő területe:
  - hálózatok struktúrájának és funkciójának megértése
  - hogyan alakulnak ki és fejlődnek
- Eddigi eredmények:
  - számos valós hálózati tulajdonságra sikerült magyarázatot találni
  - hálózatokon működő folyamatok vizsgálata (keresés, navigálás, információ szétosztás)
- A komplexitás kialakulásának okai máig sem teljesen tisztázottak
  - Számos irányított rövid kör → információfeldolgozás, kölcsönös egymásra hatás, irányítás
  - Kevés irányítási kör → jobb stabilitás

# Komplex hálózatok– Bevezetés

- Nagy hálózatok vesznek körül
- Fizikai hálók
  - Számítógép hálózatok (útvonalválasztó szint, domain szint)
  - Egyéb infrastruktúrális hálók
  - Úthálózatok
  - Ideghálózatok
  - Fehérjehálózatok
- Logikai hálók
  - Emberi kapcsolati hálózatok
  - Táplálkozási láncok
  - Metabolikus láncok
  - Bizalmi hálózatok
  - Szervezeti hálózatok
  - Genetikai hálózatok
  - WWW



# Komplex hálózatok terület célja

- A komplex hálózatok tudománya nagy hálózatok tulajdonságaival foglalkozik
  - Hogy néznek ki?
  - Milyen nagyok?
  - Milyen fő tulajdonságaik vannak?
  - Hogyan alakulnak ki és hogyan fejlődnek később?
  - Mire lehet ezeket használni?
- A terület legfontosabb sajátosságai
  - A hálózati csomópontok lokális szabályok alapján viselkednek
  - A hálózat folyamatosan alakul
  - Valós hálózatokat vizsgálunk

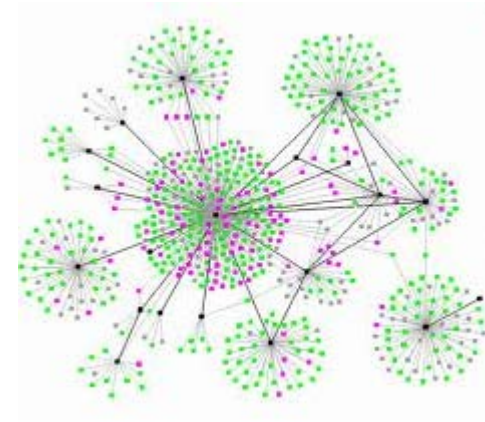
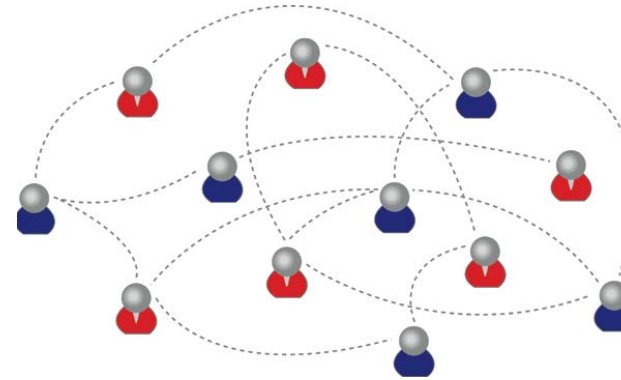


# Hálózatok kialakultása, dinamikus rendszerek

Csomópontok/ügynökök



Interakciók hálózata



- Nagy hálózatok
- Komplex dinamika: Káosz

# Kapcsolódó/alkalmazott tudományterületek

- **Biológia, orvostudomány**

- Evolúció
- Biofizika
- Genetika
- Élettan
- Idegrendszerek (anatómia, élettan)

- **Gazdaságtudomány**

- Üzleti hálózatok
- Pénzügyi hálózatok

- **Műszaki tudományok**

- Irányítás elmélet
- Dinamikus rendszerek elmélete
- Algoritmikus bonyolultságelmélet
- Információelmélet
- Statisztikus fizika

- **Szociológia**

- Kapcsolati hálók
- Csoportelmélet
- Viselkedésemélet

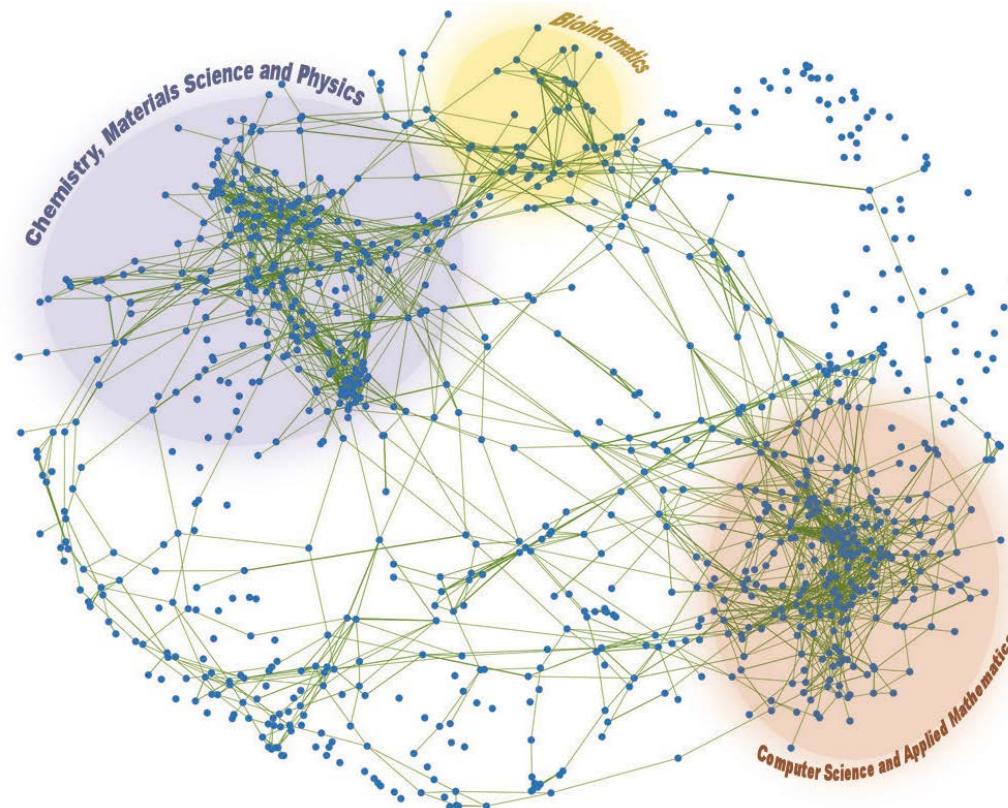
# Komplex hálózatok jelentése

- **Általános értelemben:** Nagy bonyolult hálózatok
- Hálózat komplexitása
  - Sok csomópont
  - Sok kapcsolat
  - Heterogén csomópont típusok és kapcsolattípusok
- Tisztán kivehető tendencia: kommunikációs hálózatok egyre bonyolultabbakká válnak
- Az Internet fejlődési trendek
  - Felhasználók számának drámai növekedése
  - Kicsi mobil eszközök
  - Nanotech eszközök, MEMS, szenzorok, RFID
  - Szerteágazó szabványok, sok gyártó → Heterogén eszközök
  - Virtuális hálózatok fizikai hálózatokon – VPNs, virtual ISPs
- Hogyan kezeljük ezt a komplexitást?

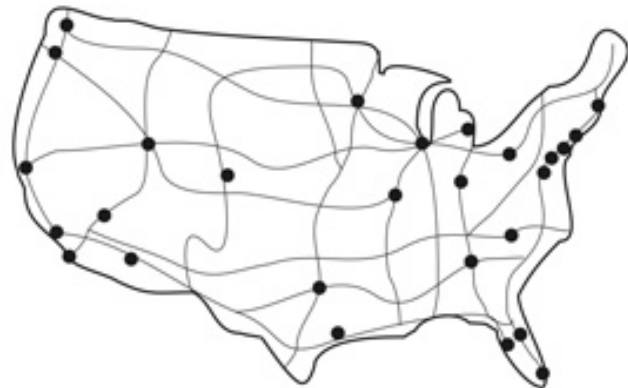
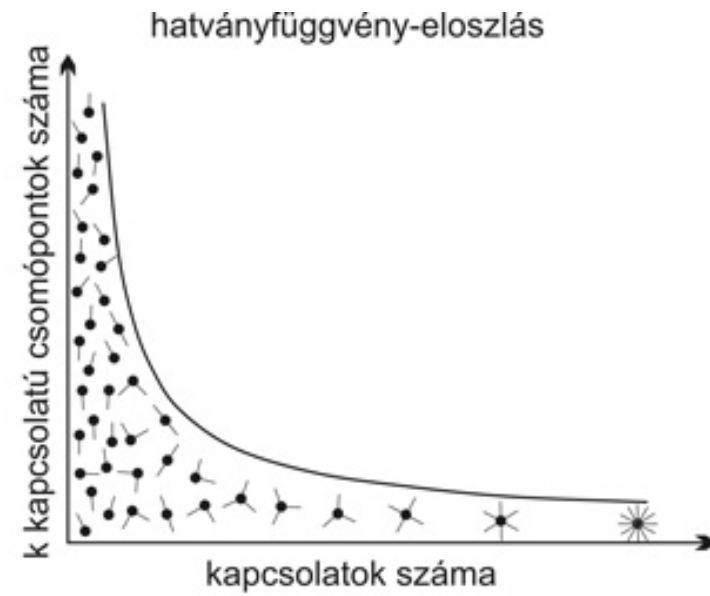
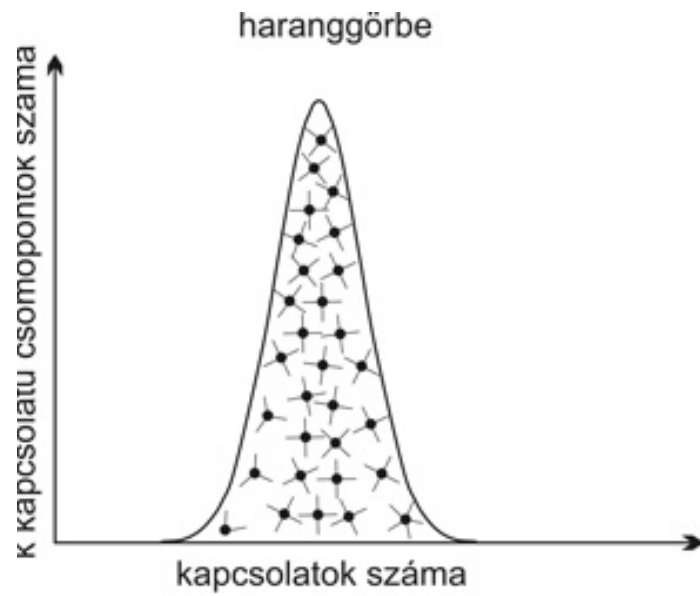


# Szűkebb értelemben vett komplex hálózatok

- Speciális értelemben:
  1. Nem teljesen véletlenszerű kapcsolatok, "csoportosuló"
  2. Kis átmérő, rövid utak, kisvilág
  3. Skálafüggetlen szerkezet: preferencia alapú kapcsolódás



# Valós hálózatok



# Valós komplex hálózat

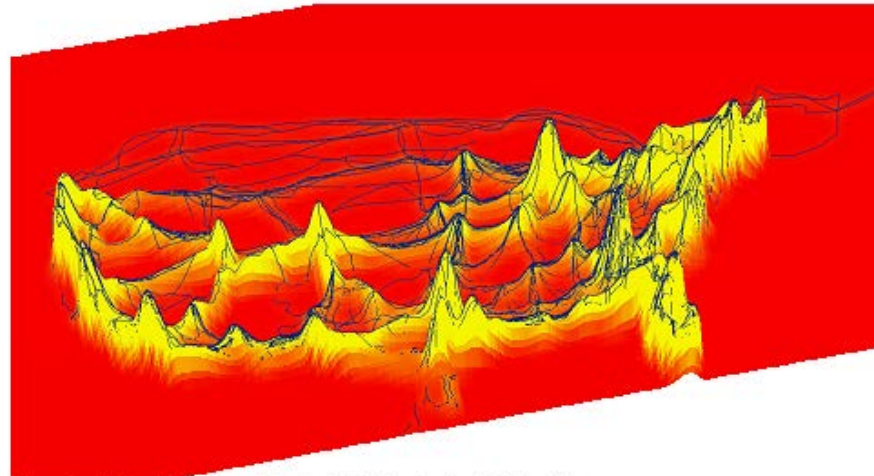


Figure 3 – The backbone network by bandwidth density for the United States

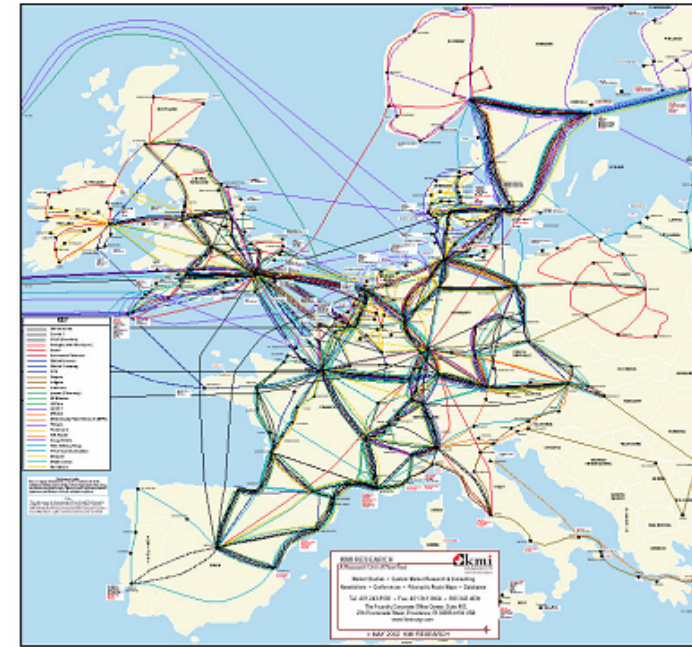
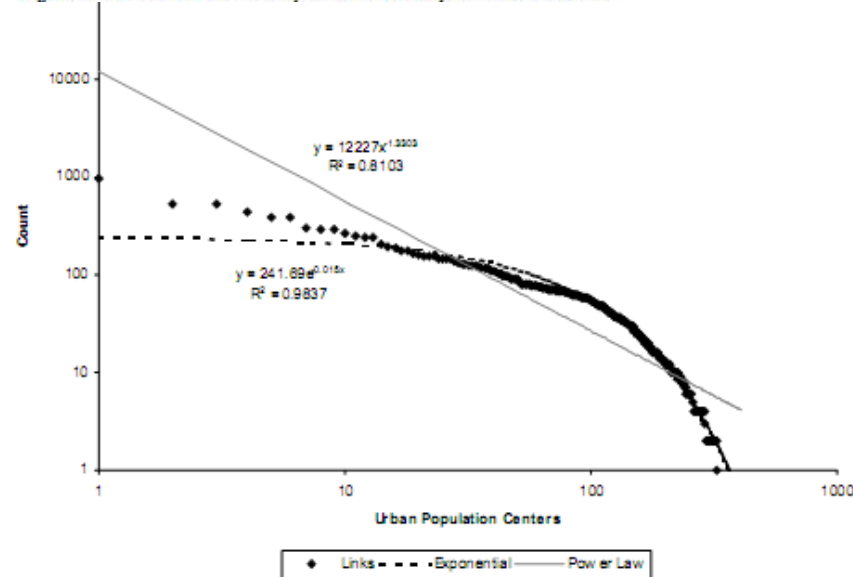
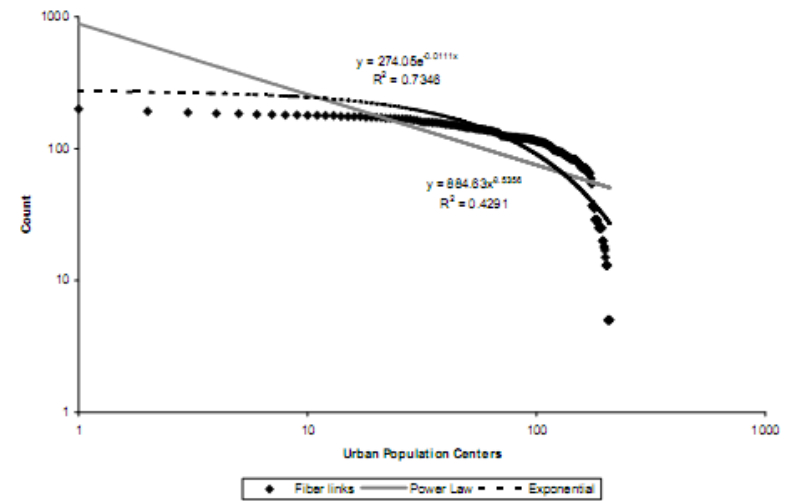
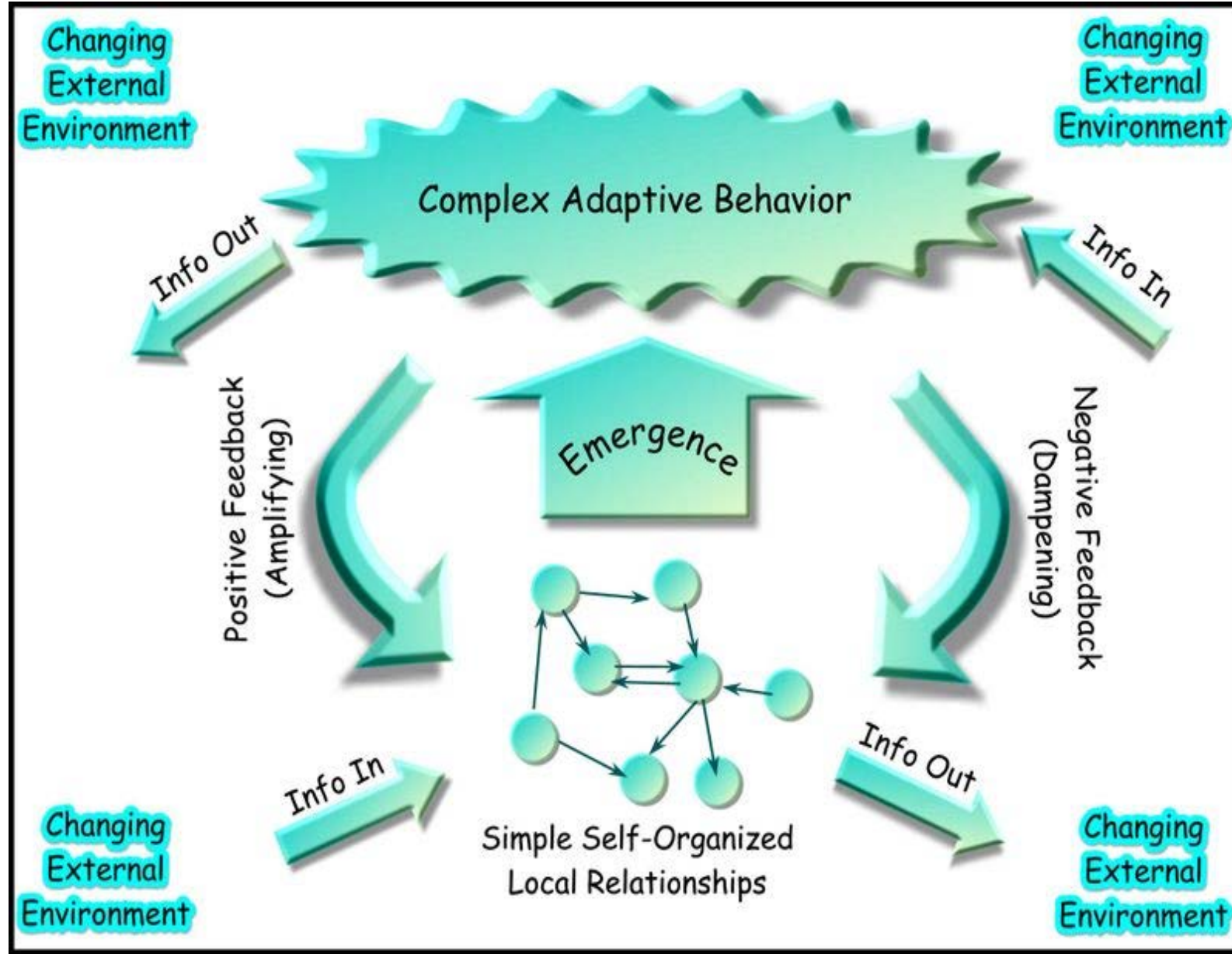


Figure 1 – Pan-European fiber optic networks (planned or in place)





# Önszerveződő rendszerek



# Game of Life

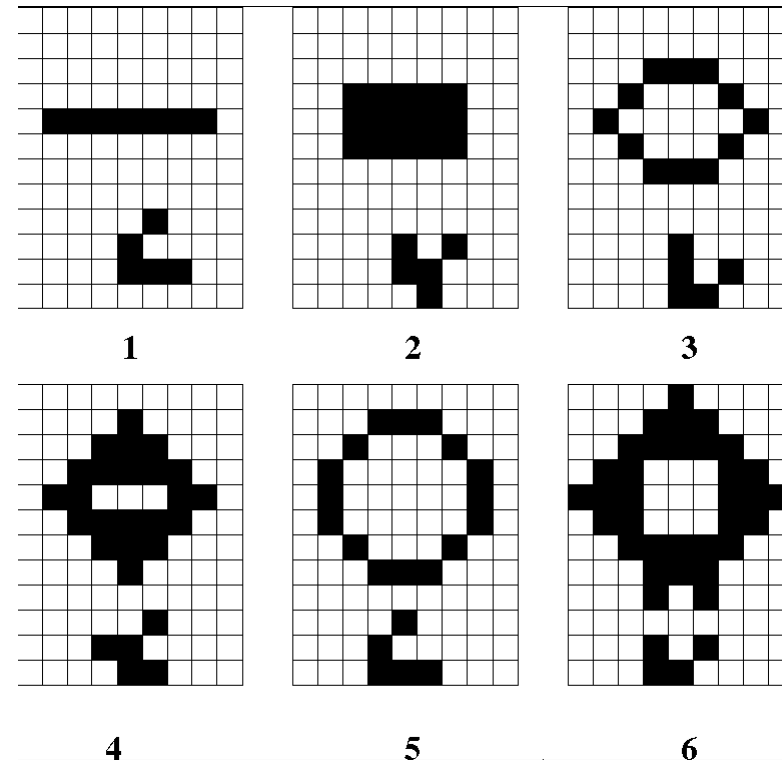
- Halott sejt három szomszédal feléled



- Sejt két vagy három szomszédal tovább él.



- Minden más esetben meghal vagy halott marad



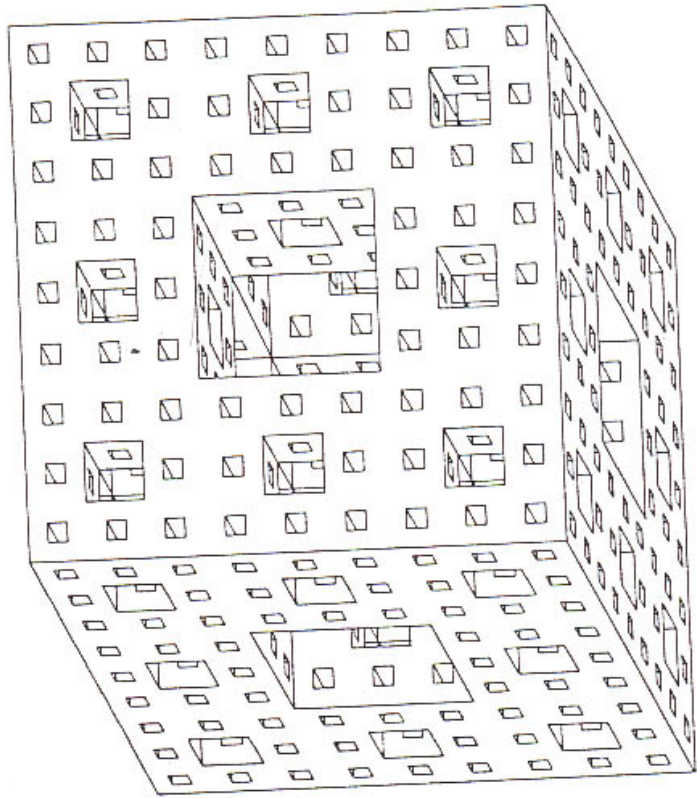
# Komplexitás és káoszelmélet

- Általában a káosz → Maximum komplexitás, teljes véletlenség
- Matematikai káosz ~ Nagy bonyolultságú rendszer
- A tudományosan vizsgált káosz a véletlentől nem függ
- Kezdeti állapot függő dinamikus rendszer (pillangó effektus)
- Fázisátmenet (pl. anyagtudomány)
- Kvantummechanika → kvantumkáosz

**Rend → Komplexitás elmélet → Rendezetlenség**

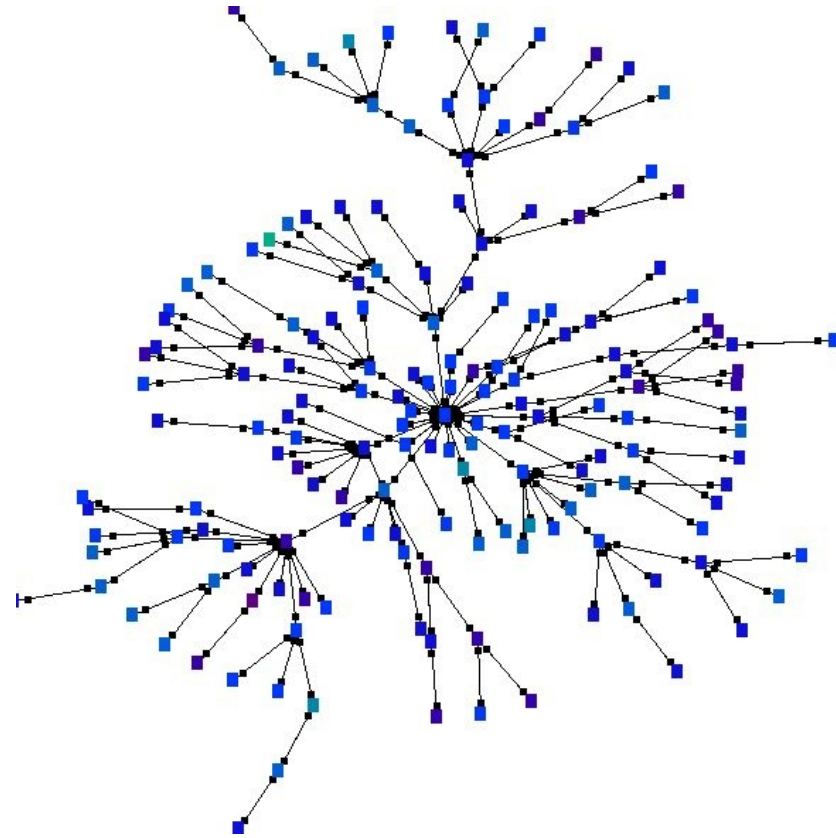
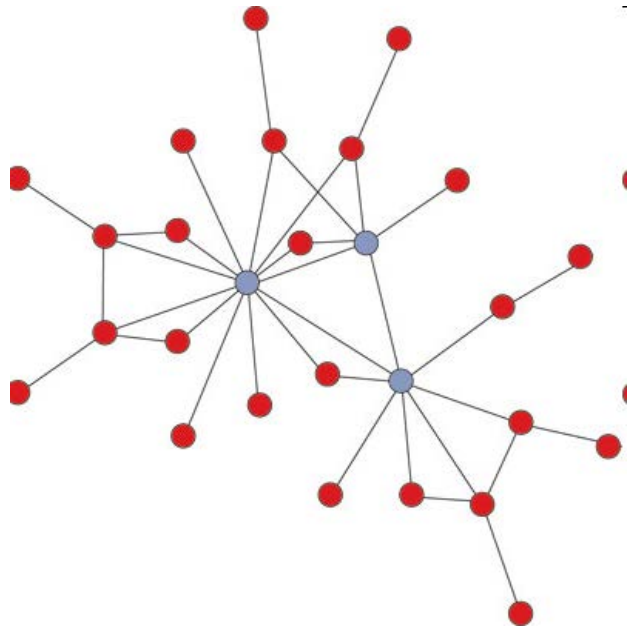
# Fraktálelmélet

- Önhasonló formák

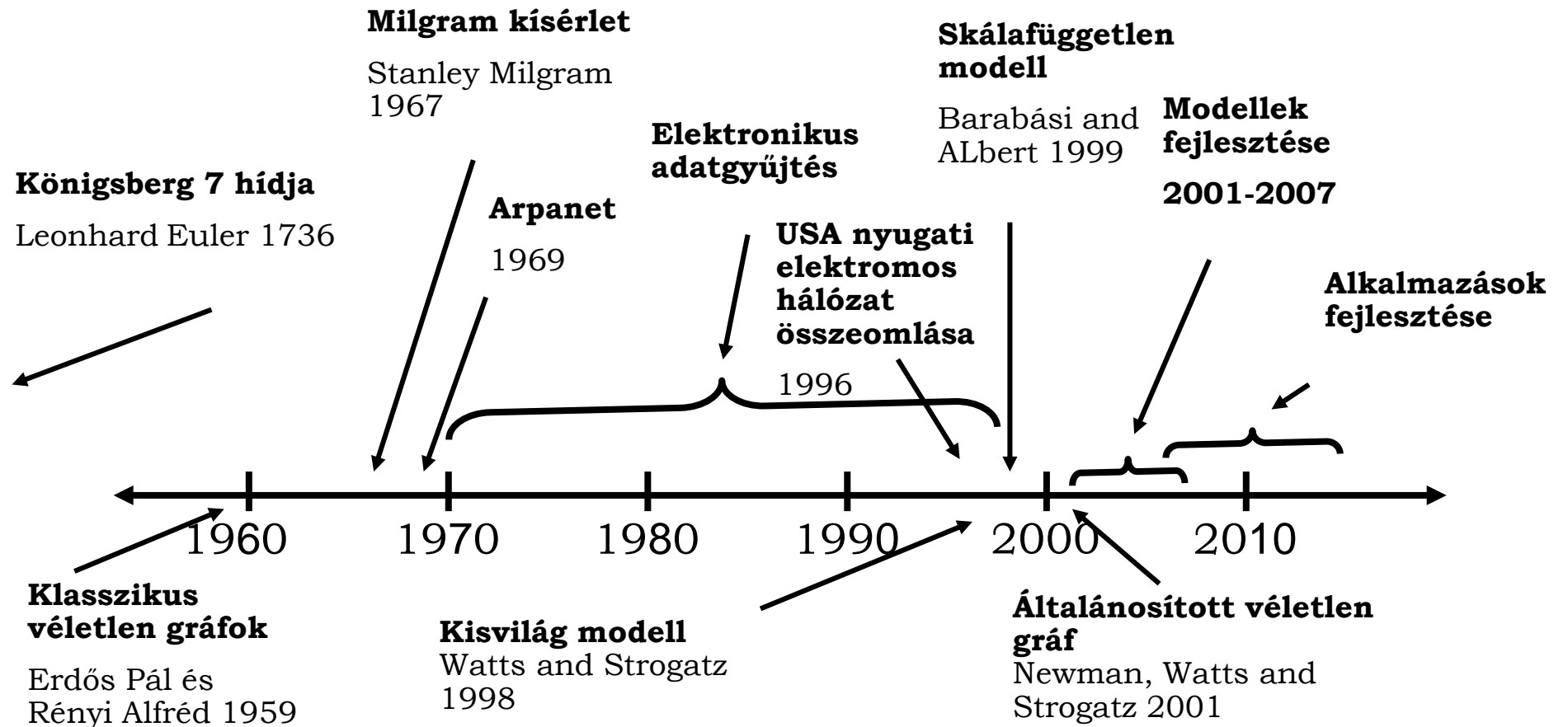




# Fraktálelmélet



# Hol tart a komplex hálózatok tudománya?



# Mi mozdította előre a nagy hálózatok vizsgálatát?

- Számítógépes adatgyűjtés
- Számítógépes gyors, automatizált feldolgozás
- Tudományterületeken átnyúló adatbázisokhoz való hozzáférés
- A tiszta redukcionista világnézet hanyatlása a tudományban
- A Internet maga nyújtotta a vizsgálat tárgyát, mint nagy hálózat!



# Pillanatsfelvételek

- A hálózatok dinamikusak
- Jelenleg nincs lehetőség a dinamizmus vizsgálatára nagy léptékben
- Legtöbb adatbázis csak a pillanatnyi állapotot tárolja
- Ezért egy-egy elemzés csak egy pillanatsfelvétel
- Előfordulnak statisztikai hibák
  - Néha később módosított eredmények

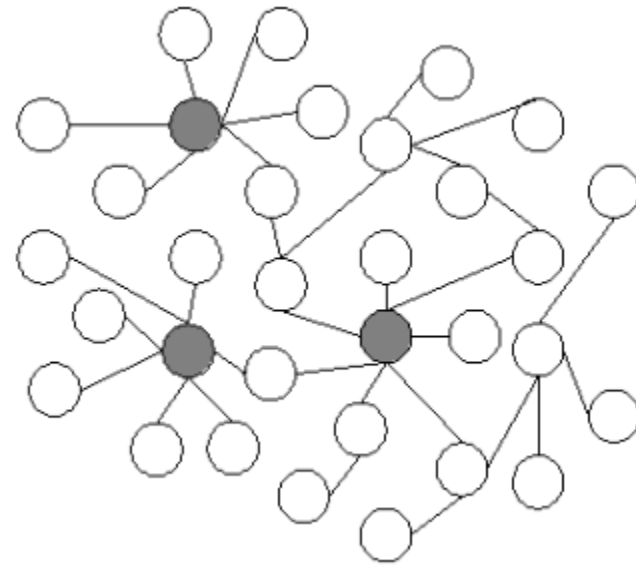
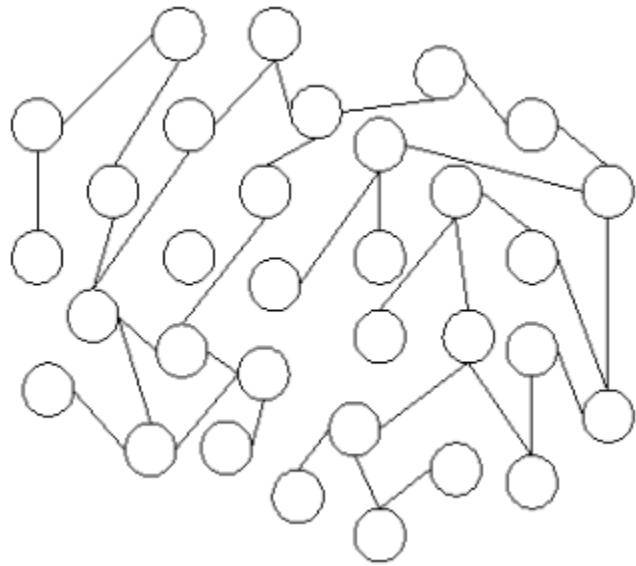




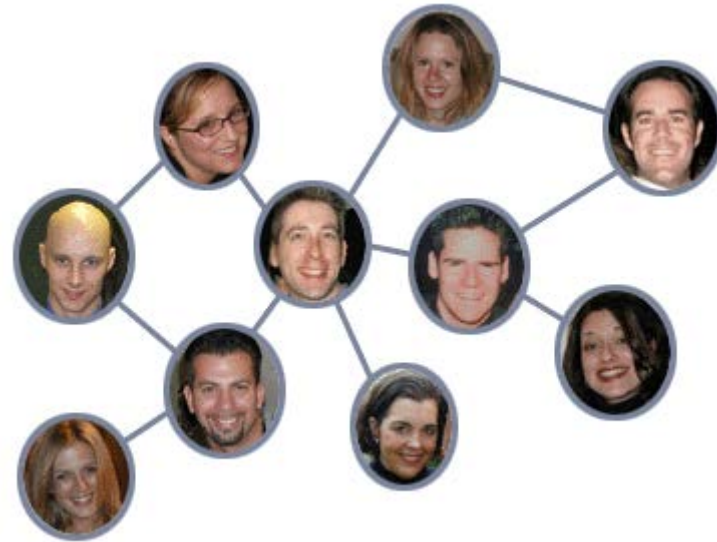
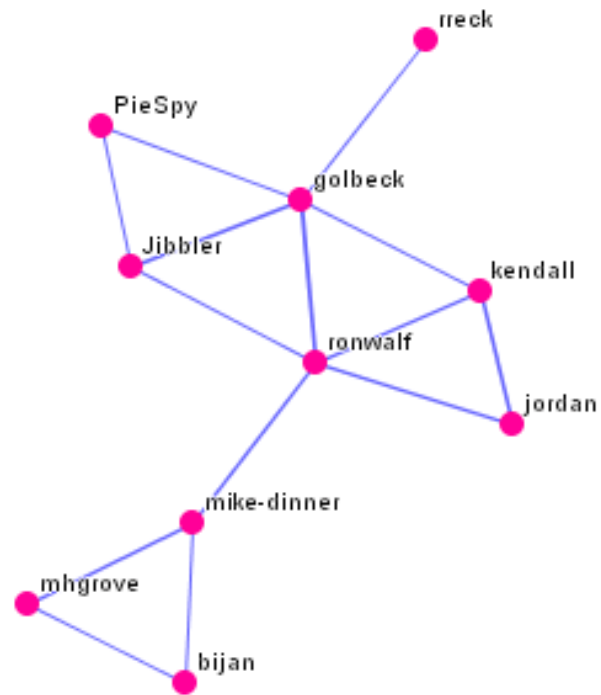


# Mérések valós hálózatokon

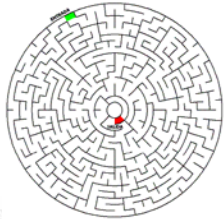
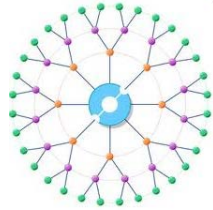
# Nagyméretű hálózatok – Van-e különbség?



# Melyik tűnik hihetőbbnek mint barátsági háló? Miért?



# Számunkra jelenleg lényeges paraméterek



## 1. Hálózat méret: Csomópontok száma

- Ezres, milliós, esetleg milliárdos méretek esetén lehet statisztikai adatokkal jól jellemezni egy hálózatot

## 2. Gruppen paraméter: “Csoportosulás” mértéke

- A barátaim jellemzően barátok-e? Ha 1 akkor mindig, ha 0 akkor soha!

## 3. Átmérő paraméter: Kis átmérő, rövid utak, “kisvilág” jelleg

- Egy rácsban igen nagy átmérők lehetnek, míg pl. a teljes gráf átmérője 1.

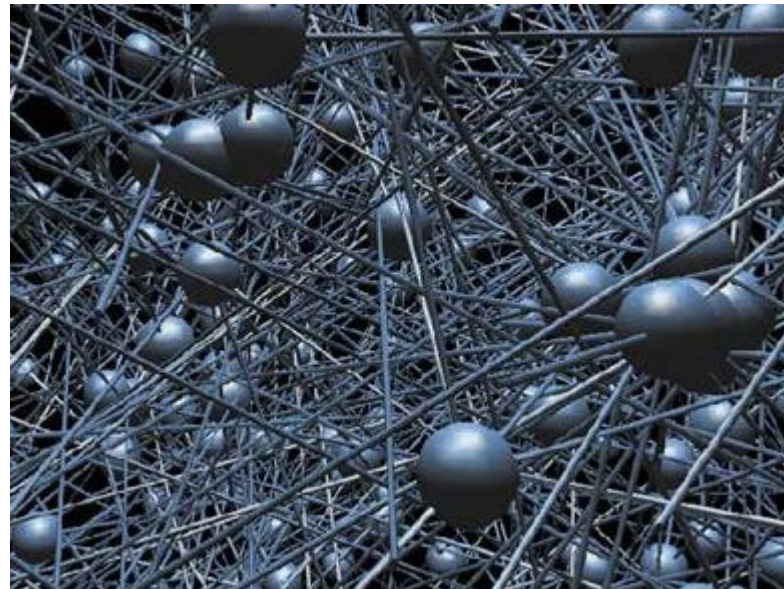
## 4. Agent Smith paraméter ( $\gamma$ ): Mennyire hasonló a szerepük? (skálafüggetlen szerkezet)

- Ha a szám magas, akkor az egyének nagyon hasonlítanak, ha alacsony akkor ( $\sim 2$ ) akkor erősen eltérő szerepek vannak



# Szociális hálózatok I.

- Csomópontok: emberek
- Kapcsolatok: kapcsolatok? Komoly probléma
- Kapcsolaterősség
- Ismeretségi háló (Dunbar-féle szám)
  - Kognitív határ átlagosan 150 fő stabil szociális kapcsolat
  - Minden egyes embert ismerek és szociológiailag viszonyítani tudom őket
  - Ennél nagyobb csoportokhoz szabályrendszer kell, törvények, politika
  - Kategóriák: közeli: 30-50, közepes: 100-200 távoli: 500-2500
- Reciprocitás (barátság hálózat)
- Módszerek:
  - Kérdőív
  - Kommunikációs intenzitás (email, telefon stb.)
  - Elégé megbízhatatlanok
- Bizonyos tulajdonságokhoz nem kell ezzel törődnünk.
  - (Focirajongók és a tea)



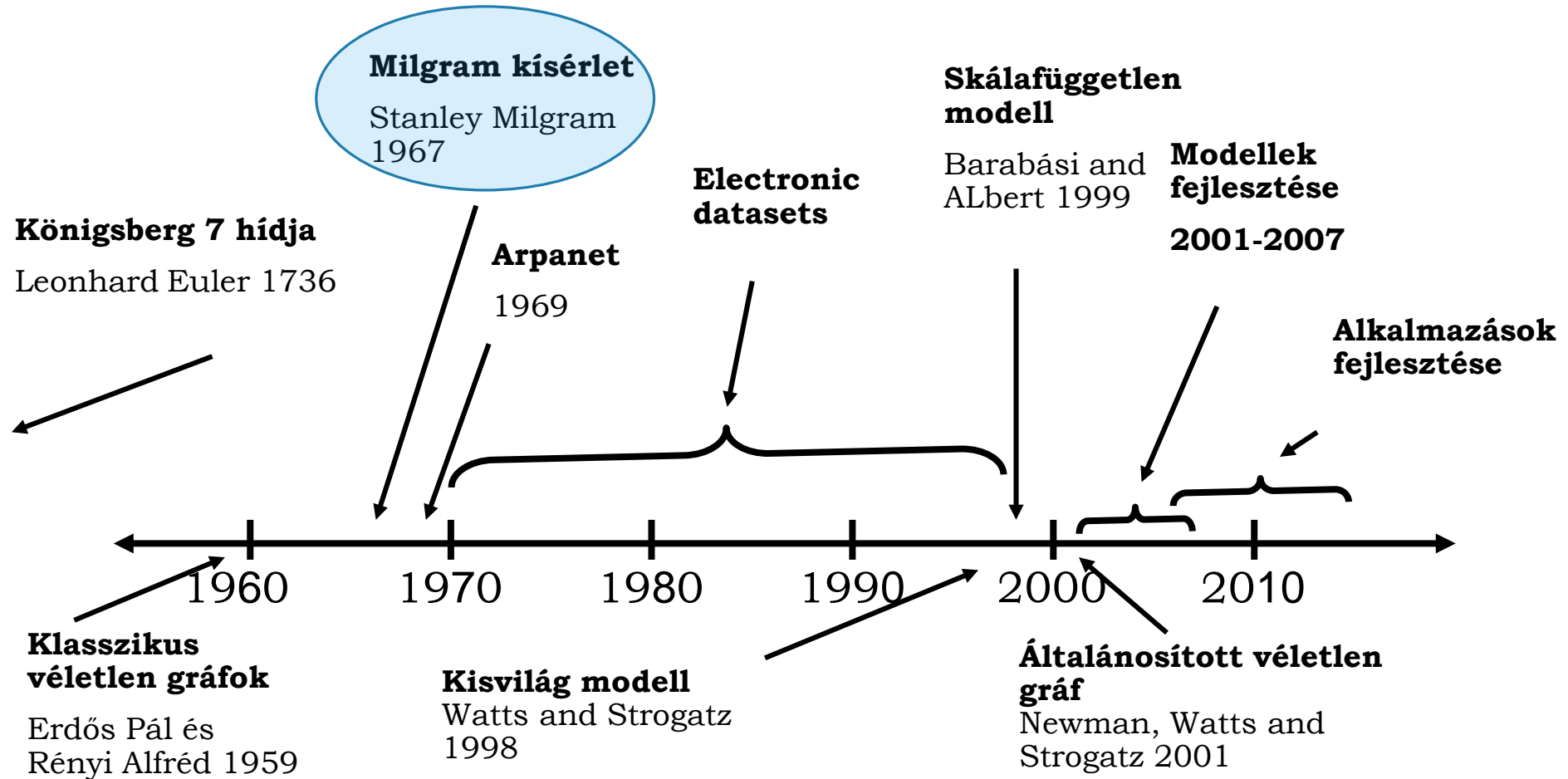
# Milgram kísérlete 1967 (a másik)

- Levélküldés: Omaha, Nebraska, Kansas → Massachusetts
- Nagy távolság (szociológiai, földrajzi)
- Véletlenszerűen választott emberek
- Információk:
  - Kísérlet célja
  - Célszemély neve, foglalkozása stb. (teológus felesége, meg egy tőzsdeügynök)
  - Szelvények
- Személyes ismeretség esetén azonnal a célhoz
- Egyébként olyanhoz aki valószínűleg személyesen ismeri+levél a Harvardra
- Eredmények:
  - Néha 1-2 lépés elég volt néha kilenc kellett
  - 296 levélből 232 nem ért célba
  - A maradékból az átlagos távolság 5.5-nek adódott (ellentmondott a tapasztalatokkal, és várakozásokkal)



1933-1984

# Kronológia





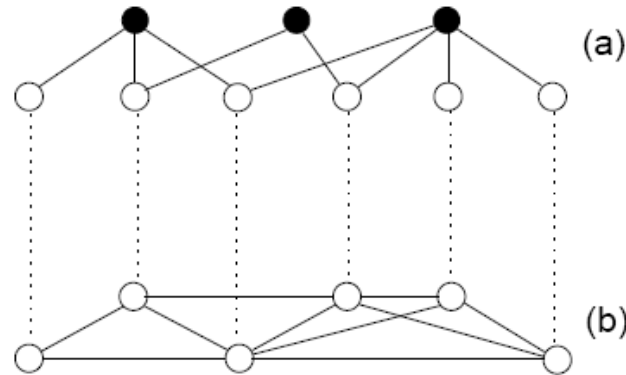


# Szociális hálózatok II.

- Központok → különleges képességgel rendelkező egyének, összekötők
  - Információ gyors eljuttatása, variálódás, változékonyság elősegítése
  - Áldozatok
  - Jockey Ewing (egy rossz példa)
- Kis foksámú pontok
  - Stabilitás
  - Állandóság
  - Fajfenntartás
  - Lehetővé teszi a központok létezését (energia)
- Futóverseny (ki miért fut)
- Központok kitermelése időbe telik
  - Diktatúra, demokrácia

# Kollaborációs hálózatok

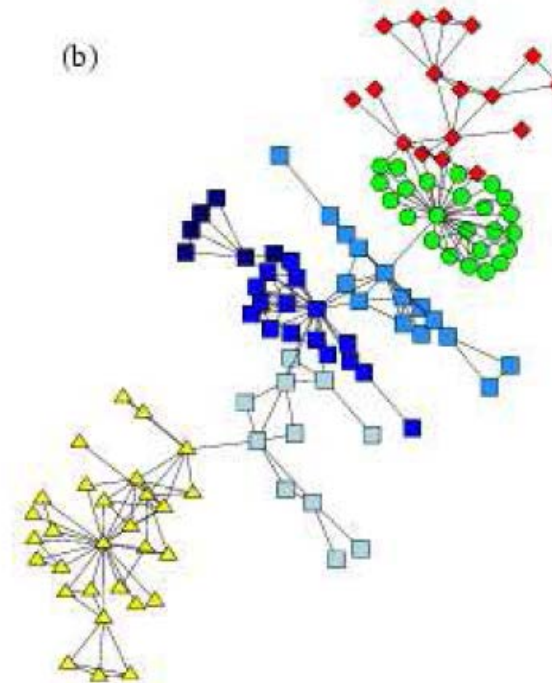
- Csomópontok: személyek és az együttműködés tárgya (tudományos cikk, film, projekt stb.)
- Kapcsolatok: adott személy részt vett-e az adott dologban
- Kétrétű hálózat:



- Vetítés: kössük össze direktbe azokat akik együttműködnek
  - Sok háromszög lesz
- Magas csoportképződési együttható

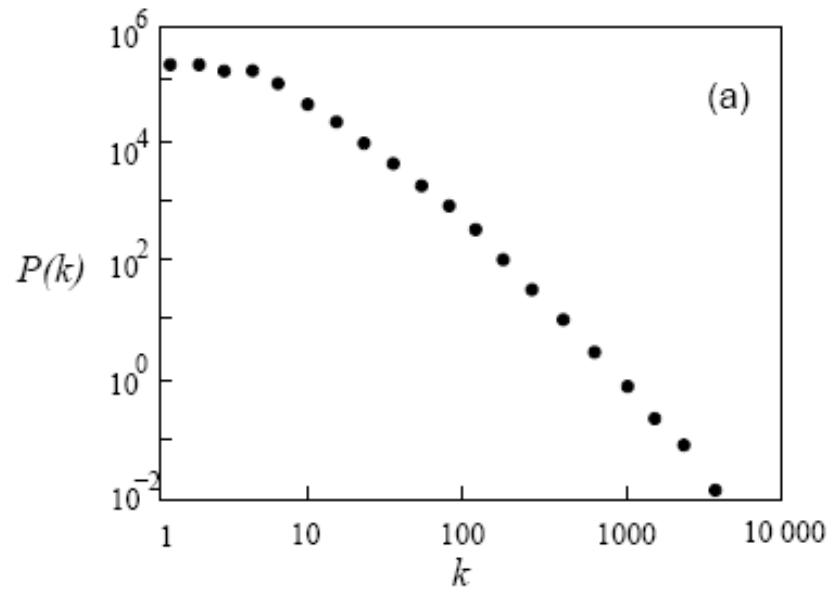
# Publikációs hálózat (Newman 2001)

- Csomópontok: személyek (fizikusok)
- Kapcsolatok: közös cikkek
- Cél: tudóstársadalom összetartása, információáramlás, tudományterületek keveredése
- Fizikusok legnagyobb adatbázisai
  - Medline, SPIRES (Stanford Public Information Retrieval System)

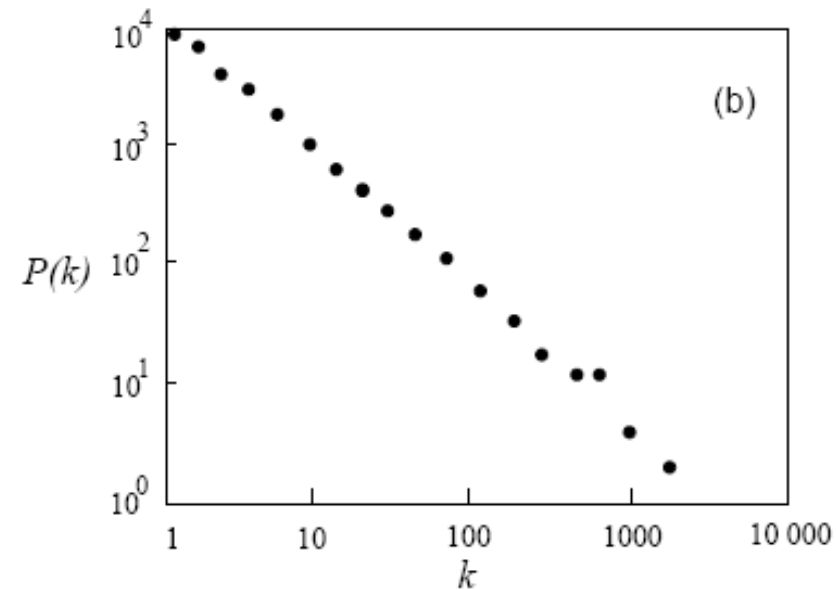


	Csomópontszám (N)	Átlagos fokszám (<k>)	Átlagos útvonal hossz <l>	Gruppen (C)	Maximális átmérő
Medline	1520251	18.1	4.6	0.066	24
SPIRES	56627	173	4.0	0.73	19

# Publikációs hálózat – kapcsolatszám eloszlás



$$\gamma \cong 3$$



$$\gamma = 1.2$$

Az egyének szerepe erősen eltér (alacsony Agent Smith paraméter)

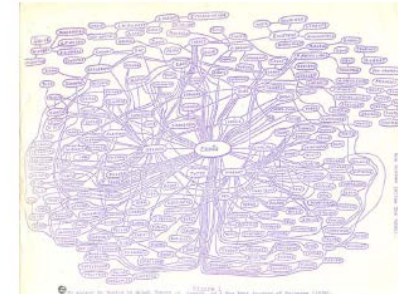
# Erdős szám <http://www.oakland.edu/enp/>

- Erdős-szám egy nemnegatív egész, amely azt mutatja, hogy az adott tudós publikálást tekintve milyen messze van Erdős Páltól
- Erdős Pál Erdős-száma 0. Egy tudós Erdős-száma  $n$ , ha az általa írt cikkek társszerzői között a legkisebb Erdős-szám  $n-1$ .
- Vagyis Erdős Pál Erdős-száma 0, valakinek az Erdős száma 1, ha írt Erdőssel közös cikket, valakinek az Erdős-száma 2, ha nem írt Erdőssel közös cikket, de írt egy 1 Erdős-számú szerzővel közösen, valakinek az Erdős száma 3, ha nem írt közös cikket sem Erdőssel, sem 1 Erdős-számúval, de írt közös cikket valamely 2 Erdős-számúval ... és így tovább.
- Más szavakkal: tekintsük a világ összes matematikai cikkeinek szerzőit egy gráf csúcsainak, és két szerzőt éllel kötünk össze, ha van olyan cikk, amelynek szerzői között mindketten szerepelnek. Ekkor Erdős-számnak nevezzük az adott személy és Erdős Pál közötti legrövidebb út hosszát ebben a gráfban.
- [Erdos1.mht](http://Erdos1.mht)

"The Erdős Number Project"  
<http://www.oakland.edu/enp>



Co-authors of Paul Erdős have number 1, co-authors of co-authors number 2, etc.



Mathematicians form a highly clustered ( $C = 0.14$ ) small world ( $L = 7.64$ )



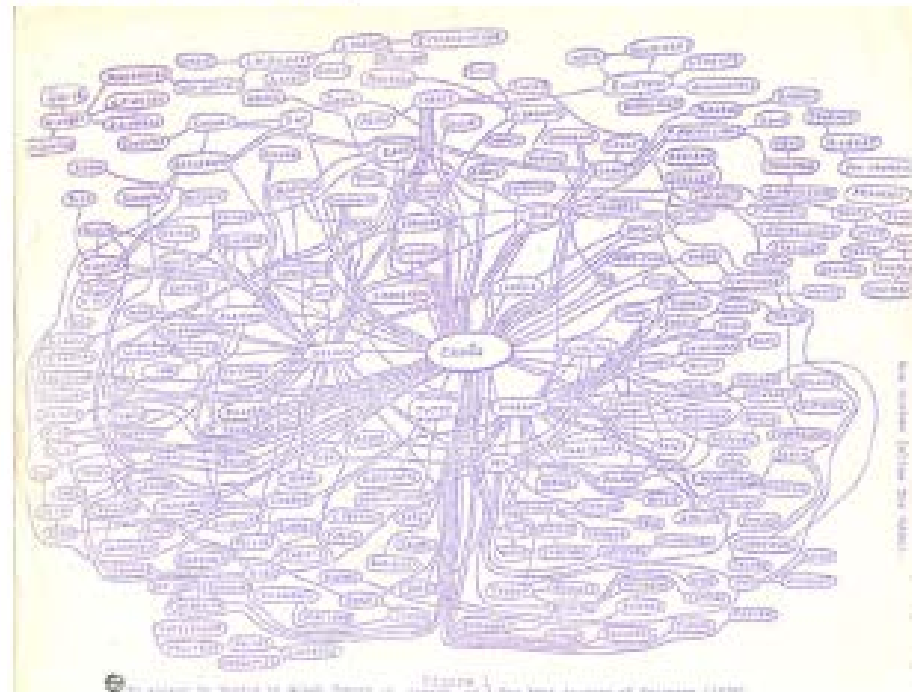
# Erdős szám

“The Erdős Number Project”

<http://www.oakland.edu/enp>



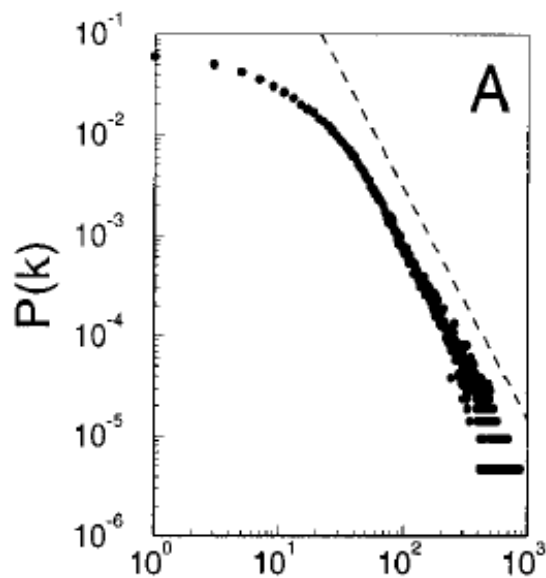
*Co-authors of Paul Erdős have number 1,  
co-authors of co-authors number 2, etc.*



*Mathematicians form a highly clustered  
( $C = 0.14$ ) small world ( $L = 7.64$ )*

# Hollywoodi színészek hálózata (Barabási 1999)

- Csomópontok: színészek
- Kapcsolatok: közös filmek
- Forrás: IMDB
- Cél: színésztársadalom összetartása → jobb filmek








Csomópont szám N	Átlagos fokszám ( $\langle k \rangle$ )	Átlagos útvonal hossz $\langle l \rangle$	Gruppen (C)
225.226	61	3.65	0.79






A fokszámeloszlás alacsony Agent Smith  
paramétert jelez:  $\gamma = 2.3$






## Bacon játék <http://oracleofbacon.org/how.php> (1994)






- Ráérős Diákok
- Weboldal: Wasson és Tjaden 1997-ben a Time által kiválasztott 10 legnépszerűbb oldal közé
- IMDB-ről letöltött adatbázis
- Szolgáltatások:
  - Kapcsolat meghatározása két színész között
  - Színész központiséga
- Bacon központiséga [1000kozpont](#)
  - 0 1
  - 1 1806
  - 2 145024
  - 3 395126
  - 4 95497
  - 5 7451
  - 6 933
  - 7 106
  - 8 13






# Bacon játék






**KEVIN KLINE, BA'70** starred in *Wild, Wild, West* (1999) with  Will Smith who was in *Bad Boys* (1995) with  who was in *Family Man* (2000) with  Nicolas Cage who was in *Honeymoon in Vegas* (1992) with  who was in *Footloose* (1984) with the shining star  KEVIN BACON

**IU School of Medicine's LAWRENCE H. EINHORN, BA '85** headed the cancer treatment team of  Lance Armstrong who was in *Dodgeball* (2004) with  who was in the *Psycho* (1998) remake with  Anne Heche who was in *Volcano* (1997) with  who was in *JFK* (1991) with the man known as  KEVIN BACON

**JOHN MELLENCAMP, DM HON '00** starred in *Falling From Grace* (1992) with  who was in *The Contender* (2000) with  Joan Allen who was in *The Bourne Supremacy* (2004) with  Matt Damon who was in *Ocean's 11* (2001) with  who was in *Flatliners* (1990) with the one, true  KEVIN BACON

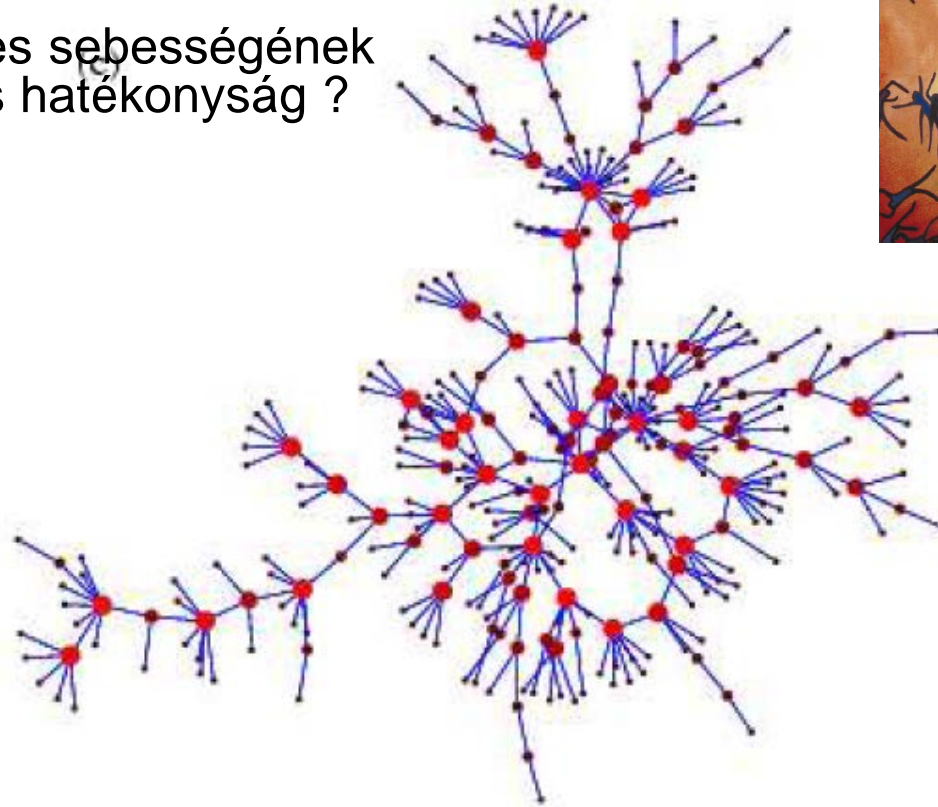
**JOSHUA BELL, ArtD '89** was in *Music of the Heart* (1999) with  Kieran Culkin who was in *Cider House Rules* (1999) with  who was in *Wonder Boys* (2000) with  Michael Douglas who was in *The Game* (1997) with  who was in *Mystic River* (2003) with the charming  KEVIN BACON

**TAVIS SMILEY, BS'03, LHD'04** guest starred on the TV series *American Dreams* (2004) with  Gail O'Grady who was on the *NYPD Blue* (1993) TV series with  who was in *Star Wars: Attack of the Clones* (2002) with  Natalie Portman who was in *Mars Attacks* (1996) with  who was in *A Few Good Men* (1992) with the irresistible  KEVIN BACON

**IUB Slavic Languages faculty member LAURENCE RICHTER** is the father of TV star  Andy Richter who was in *Dr. T and the Women* (2000) with  who was in *First Knight* (1995) with  Sean Connery who was in *The Rock* (1996) with  who was in *Apollo 13* (1995) with the center of the universe  KEVIN BACON

# Szexuális kapcsolati háló (Liljeros 2001)

- Csomópontok: emberek
- Kapcsolatok: ki-kivel-mikor
- Ilyet nem mindenhol lehet megcsinálni
- Cél: génkicserélődés sebességének növelése, evolúciós hatékonyság ?

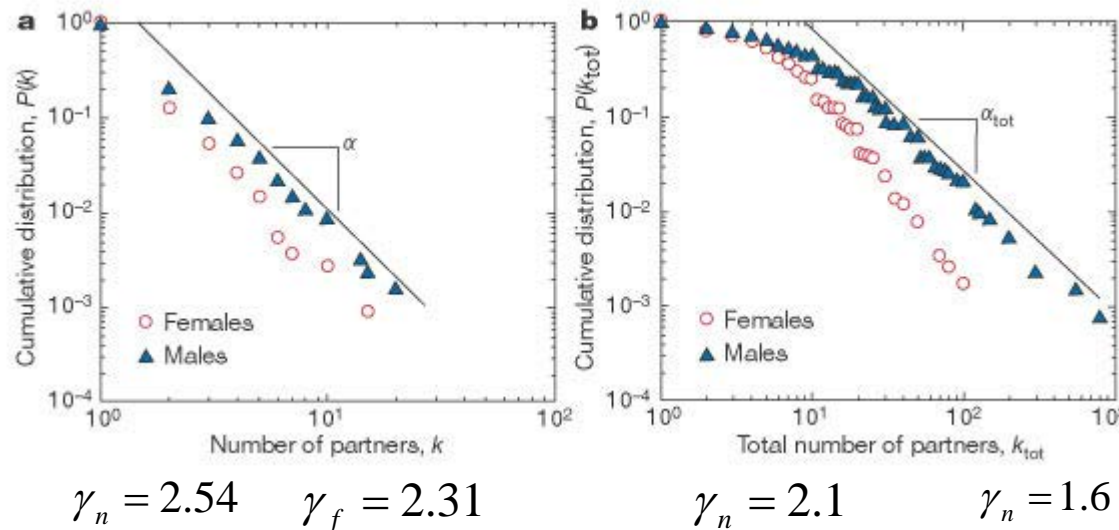


- Bővebb infó: <http://www2.sociology.su.se/home/Liljeros/Nature.pdf>



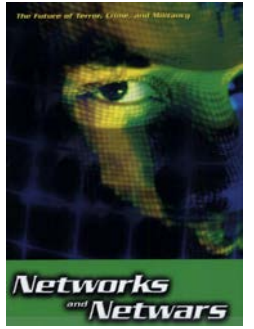
# Szexuális kapcsolati háló

- 4,781 18–74 éves svéd férfi és nő interjúval és kérdőívvel
- 59%-os válaszarány, 2,810 csomópontból álló reprezentatív minta
- 1. Szexuális kapcsolatok száma 12 hónappal a felmérés előtt
- A férfiak nagyobb partnerszámot jelentettek be mint a nők (szociális elvárások)
- 2. Egés



Nem látható tipikus jelleg, nincs “meghatározó viselkedés”.

# Terrorista hálózat (Krebs 2002)



- Problémák
  - Nem teljes információ
  - Nem egyértelmű, hogy ki tartozik bele és ki nem
  - Dinamikus
  - A rejtett hálózatok elemzése egészen más módszereket kíván
  - Nehezen levonható következtetések
- 19 gépeltérítő
- 3 fajta kapcsolaterősség (együtt töltött idő alapján)
  - Közvetlen: egy iskolába jártak stb.
  - Lazább: együtt utaztak, közös találkozók
  - Távoli: egyszeri közös ügyletek



# Terrorista hálózat

## THE HIJACKERS ...

### American Airlines 11

Crashed into WTC (north)



**Mohamed Atta**  
(Egyptian)  
Received pilot training



**Waleed M. Alshehri**  
(Saudi)  
Commercial pilot



**Wail Alshahri**  
(Saudi)  
Possible pilot training



**Satam al-Suqami**  
(Nationality unknown)



**Abdulaziz Alomari\***  
(Saudi)  
Possible pilot training

### American Airlines 77

Crashed into Pentagon



**Khalid al-Midhar**  
(Nationality unknown)  
Received pilot training



**Majed Moqed**  
(Nationality unknown)



**Salem Alhamzi\***  
(Saudi)  
Possible pilot training



**Nawaf Alhamzi\***  
(Saudi)

### United Airlines 175

Crashed into WTC (south)



**Marwan al-Shehhi**  
(United Arab Emirates)  
Received pilot training



**Fayez Ahmed**  
(Believed to be Saudi)



**Ahmed Alghamdi**  
(Possibly Saudi)



**Hamza Alghamdi**  
(Believed to be Saudi)  
Possible pilot training



**Mohald Alshehri**  
(Nationality unknown)  
Possible pilot training

### United Airlines 93

Crashed in Pennsylvania



**Ziad Jarrah**  
(Lebanese)  
Received pilot training



**Ahmed Alhaznawi**  
(Saudi)



**Ahmed Alnami**  
(Nationality unknown)



**Saeed Alghamdi\***  
(Seems to be Saudi)

## AND HOW THEY WERE CONNECTED

### Attended same technical college

Hamburg, Germany

Mohamed Atta  
Marwan al-Shehhi  
Ziad Jarrah

### Took flight classes together

Pilot schools  
in Florida

Mohamed Atta  
Marwan al-Shehhi

Pilot schools  
in San Diego

Khalid al-Midhar  
Nawaf Alhamzi

### Known to be together in week before attacks

Stayed together  
in a Florida  
motel

Mohamed Atta  
Marwan al-Shehhi

Attended a gym  
in Maryland  
(Sept 2-6),  
also seen dining  
together

Khalid al-Midhar  
Majed Moqed  
Salem Alhamzi  
Nawaf Alhamzi  
Hani Hanjour

### Last known address

Hollywood, Florida

Marwan al-Shehhi  
Waleed M. Alshehri  
Wail Alshahri  
Ziad Jarrah  
Hani Hanjour

Other cities  
in Florida

Mohamed Atta  
Fayez Ahmed  
Ahmed Alghamdi  
Mohald Alshehri  
Khalid al-Midhar  
Ahmed Alhaznawi  
Ahmed Alnami  
Saeed Alghamdi

### Bought flight tickets using same address

• Mohamed Atta\*  
Marwan al-Shehhi  
Abdulaziz Alomari\*

\* Also used same credit card

• Waleed M. Alshehri  
Wail Alshahri

• Fayez Ahmed  
Mohald Alshehri

### Bought flight tickets together

Mohamed Atta  
Ziad Jarrah  
Ahmed Alhaznawi

Picked up tickets  
bought earlier in  
Baltimore

Khalid al-Midhar  
Majed Moqed

Bought from the  
same travel agent

Outside Florida

Satam al-Suqami  
Hamza Alghamdi  
Abdulaziz Alomari  
Majed Moqed  
Salem Alhamzi  
Nawaf Alhamzi



# Terrorista hálózat

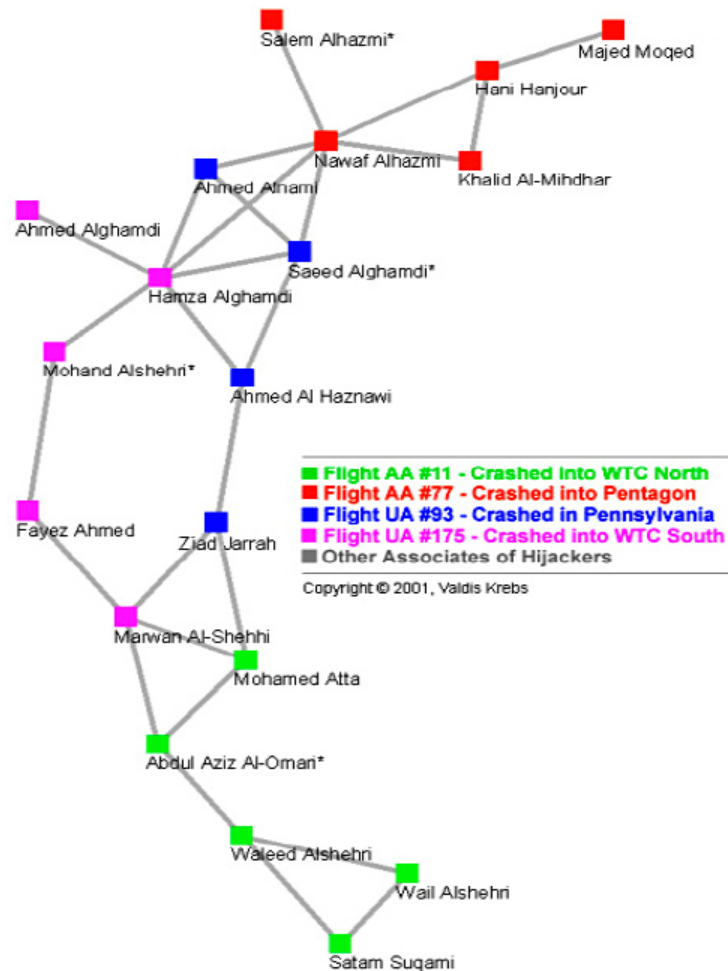


Figure 2 Trusted Prior Contacts

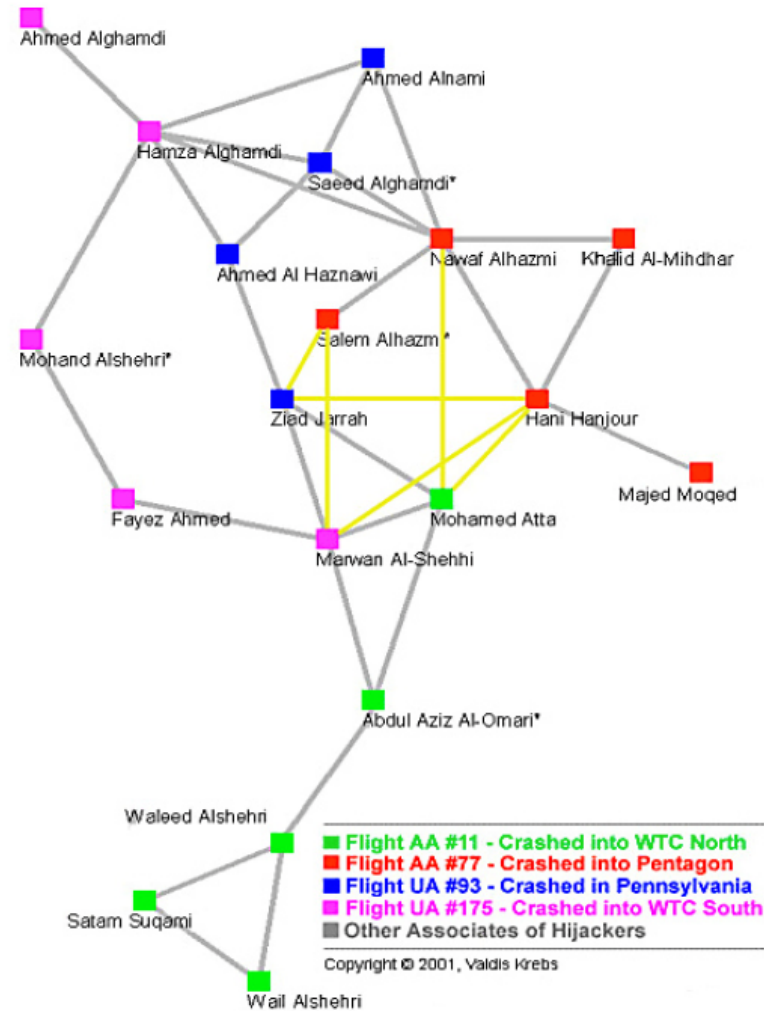


Figure 3 Trusted Prior Contacts + Meeting Ties [shortcuts]



# Terrorista hálózat

**Table 1. Without shortcuts**

Name	Cluster- ing Coef- ficient	Mean Path Length	Short- cuts
Satam Suqami	1.00	5.22	0.00
Wail Alshehri	1.00	5.22	0.00
Majed Moqed	0.00	4.67	0.00
Waleed Alshehri	0.33	4.33	0.33
Salem Alhazmi*	0.00	3.89	0.00
Khalid Al-Mihdhar	1.00	3.78	0.00
Hani Hanjour	0.33	3.72	0.00
Abdul Aziz Al-Omari*	0.33	3.61	0.33
Ahmed Alghamdi	0.00	3.50	0.00
Ahmed Alnami	1.00	3.17	0.00
Mohamed Atta	0.67	3.17	0.00
Marwan Al-Shehhi	0.33	3.06	0.25
Fayez Ahmed	0.00	2.94	1.00
Nawaf Alhazmi	0.27	2.94	0.00
Ziad Jarrah	0.33	2.83	0.33
Mohand Alshehri*	0.00	2.78	1.00
Saeed Alghamdi*	0.67	2.72	0.00
Ahmed Al Haznawi	0.33	2.67	0.33
Hamza Alghamdi	0.27	2.56	0.17
<b>Overall</b>	<b>0.41</b>	<b>4.75</b>	<b>0.19</b>

**Table 2. With shortcuts**

Name	Cluster- ing Coef- ficient	Mean Path Length	Short- cuts
Satam Suqami	1.00	3.94	0.00
Wail Alshehri	1.00	3.94	0.00
Ahmed Alghamdi	0.00	3.22	0.00
Waleed Alshehri	0.33	3.06	0.33
Majed Moqed	0.00	3.00	0.00
Mohand Alshehri*	0.00	2.78	1.00
Khalid Al-Mihdhar	1.00	2.61	0.00
Ahmed Alnami	1.00	2.56	0.00
Fayez Ahmed	0.00	2.56	1.00
Ahmed Al Haznawi	0.33	2.50	0.33
Saeed Alghamdi*	0.67	2.44	0.00
AbdulAziz Al-Omari*	0.33	2.33	0.33
Hamza Alghamdi	0.27	2.28	0.17
Salem Alhazmi*	0.33	2.28	0.33
Ziad Jarrah	0.40	2.17	0.20
Marwan Al-Shehhi	0.33	2.06	0.17
Hani Hanjour	0.33	2.06	0.00
Mohamed Atta	0.50	1.94	0.00
Nawaf Alhazmi	0.24	1.94	0.14
<b>Overall</b>	<b>0.42</b>	<b>2.79</b>	<b>0.18</b>

\* suspected to have false identification



# Terrorista hálózat

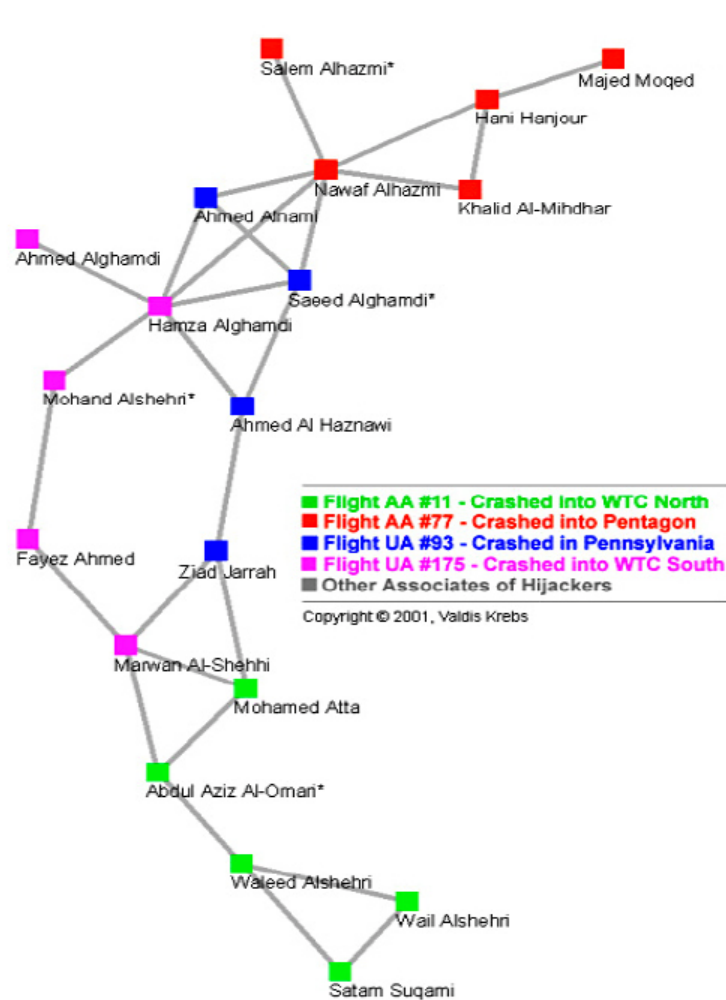
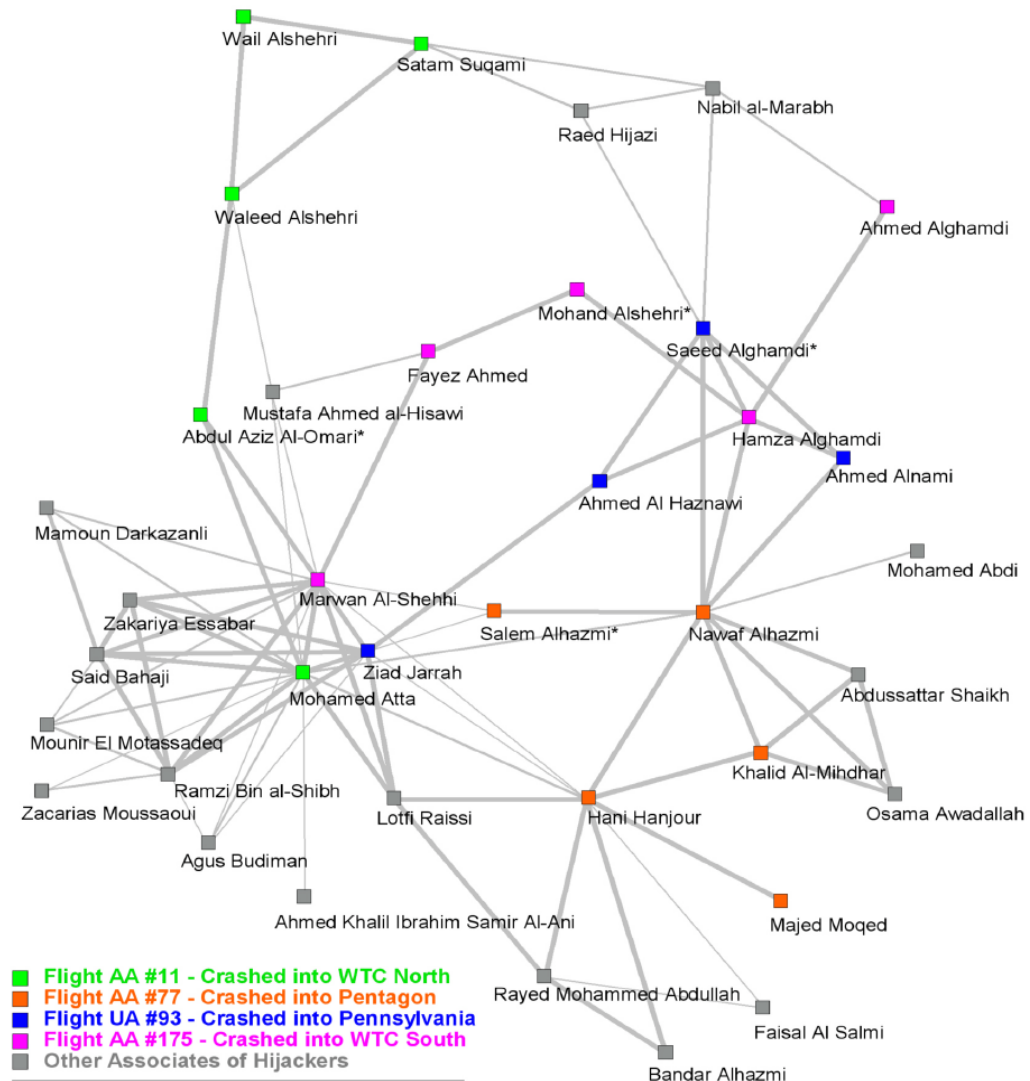


Figure 2 Trusted Prior Contacts



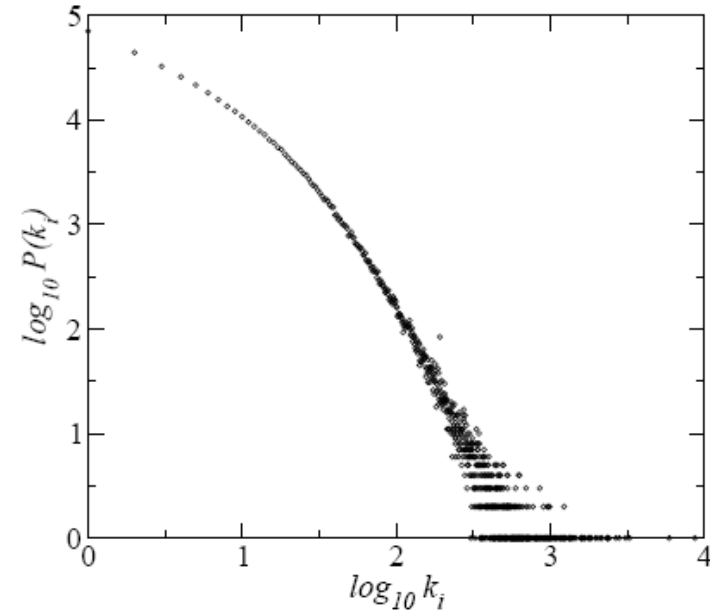
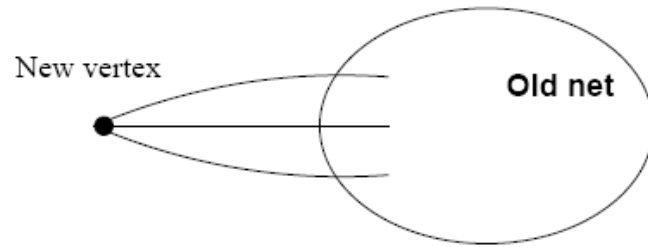
További infók: <http://www.orgnet.com/hijackers.html>

# Hivatkozások hálózata (Redner 1998)

- Csomópontok: cikkek
- Kapcsolatok: hivatkozott cikkek
- Irányított gráf
- Adatok: ISI (Institute for Scientific Information) adatbázis és Physical Review
- Cél: Információ stuktúrált eljuttatása az olvasóhoz, tudományszervezés
- Gazdag még gazdagabb lesz jelenség

N	$\langle k_i \rangle$	$\langle l \rangle$	C
783.339	8.57	-	-

# Hivatkozások hálózata



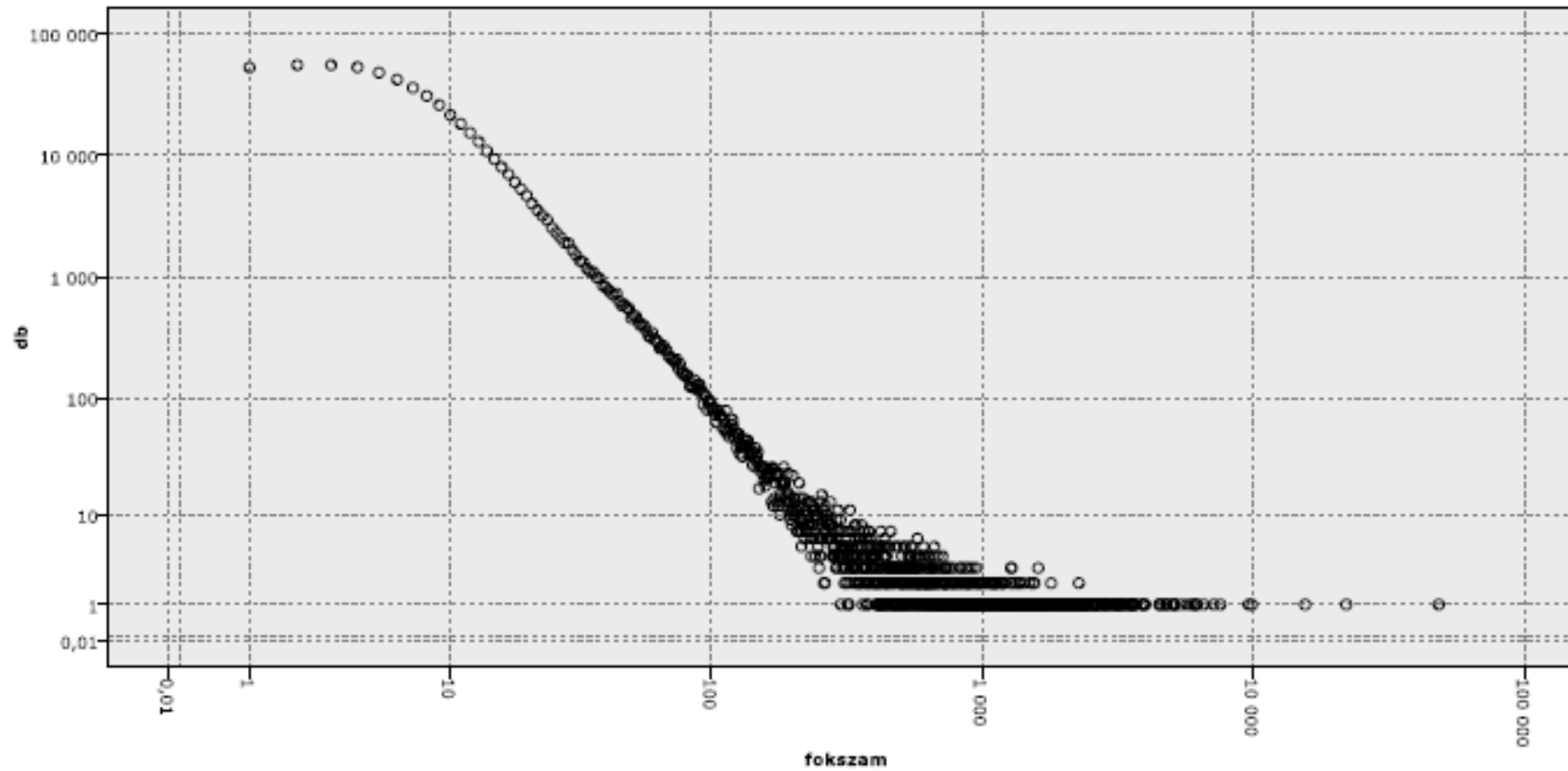
$$\gamma = 3 \quad \gamma = 2.9 \quad \gamma = 2.5$$

# Telefonhívások (Szabó-Gulyás-Heszberger 2009)

- Nincs sok adat (titkossági okok)
- Csomópontok: telefonszámok
- Kapcsolatok: hívások
- Irányított gráf
- Budapest 643135

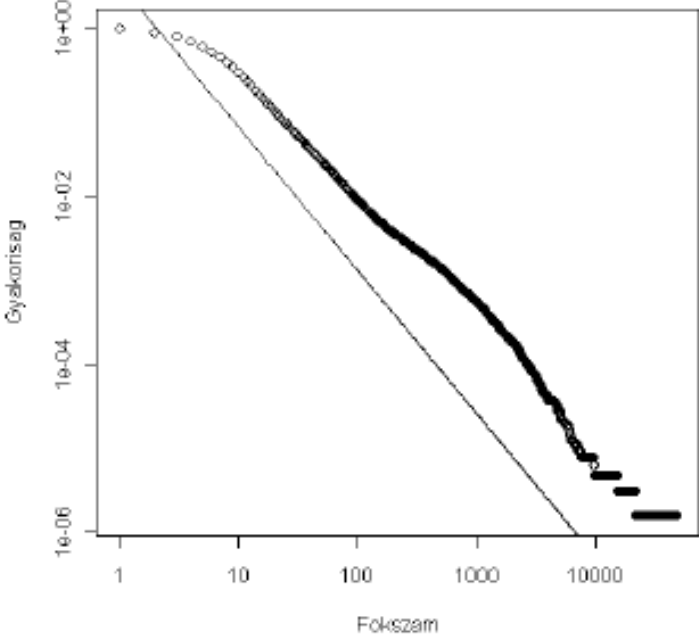
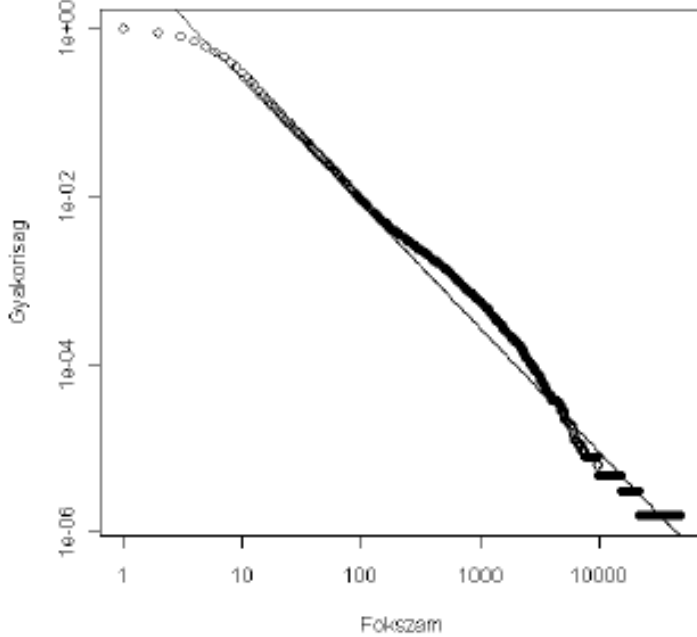
N	$\langle K \rangle$	$\langle l \rangle$	C_1	C_2	C_v	C_1/C_V	C_2/C_v
643135	12.48	4.86	0.0742	0.0043	1.9E-5	3825.8	221.9

# Telefonhívások

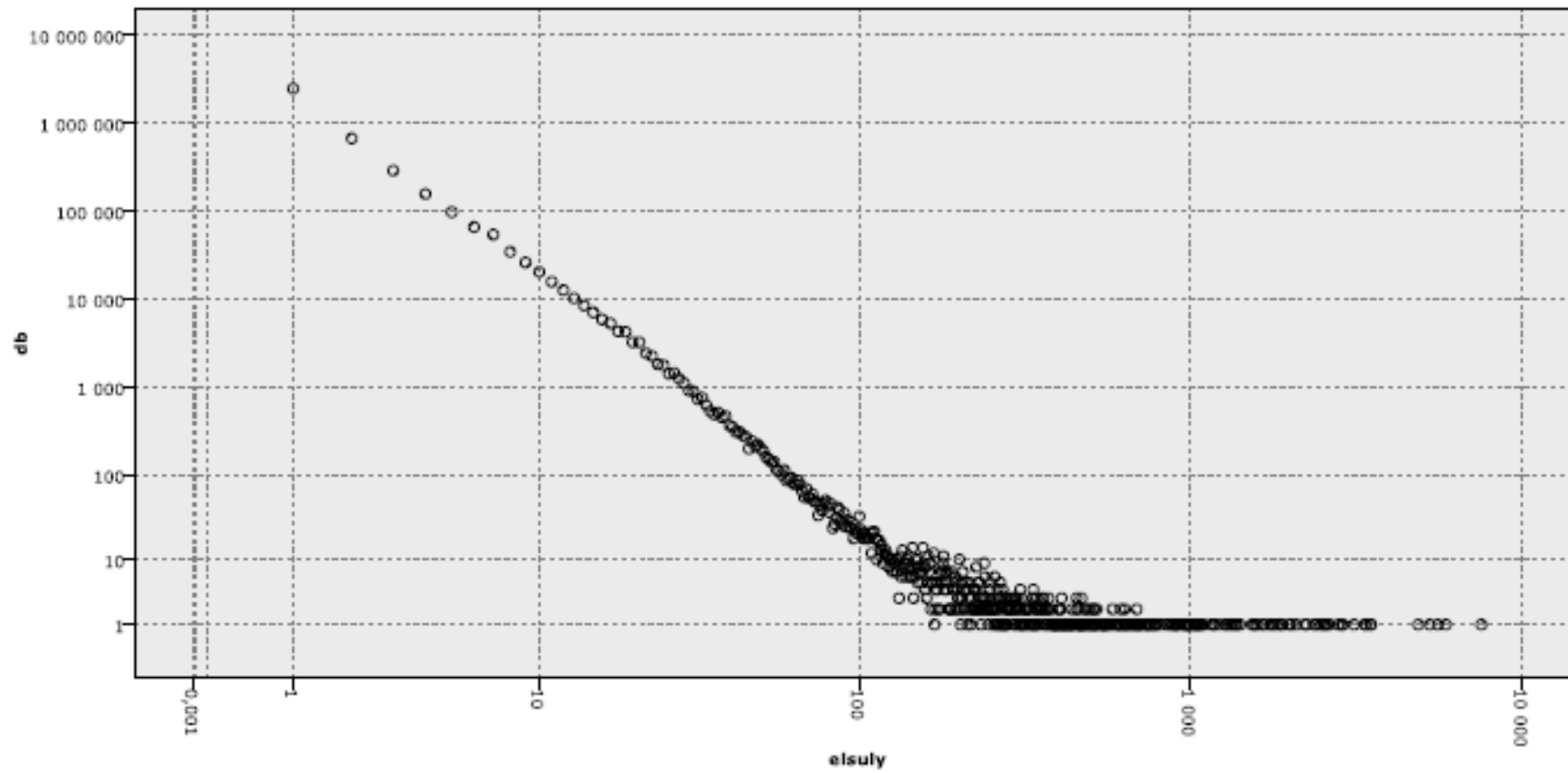




# Telefonhívások

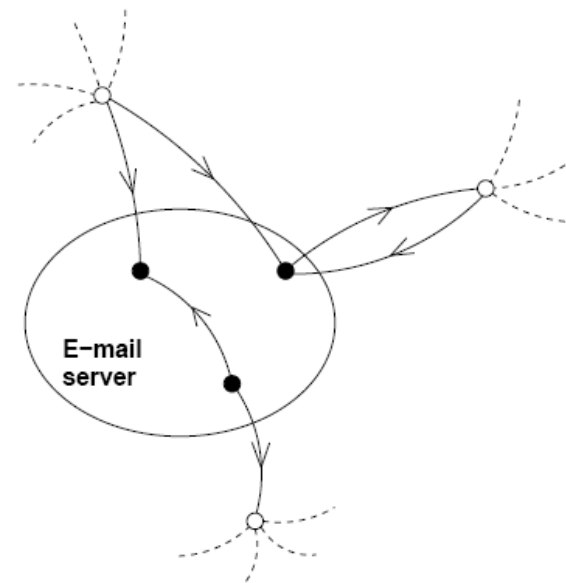
	Illesztés newman módszerrel:	Illesztés lineáris regresszióval
	<p>The degree distribution in a log-log scale</p>  <p>Gyakoriság</p> <p>Fokszám</p>	<p>The degree distribution in a log-log scale</p>  <p>Gyakoriság</p> <p>Fokszám</p>
$\gamma$	2.709519	2.481185
$\gamma(\text{cum})$	1.709519	1.481185

# Telefonhívások

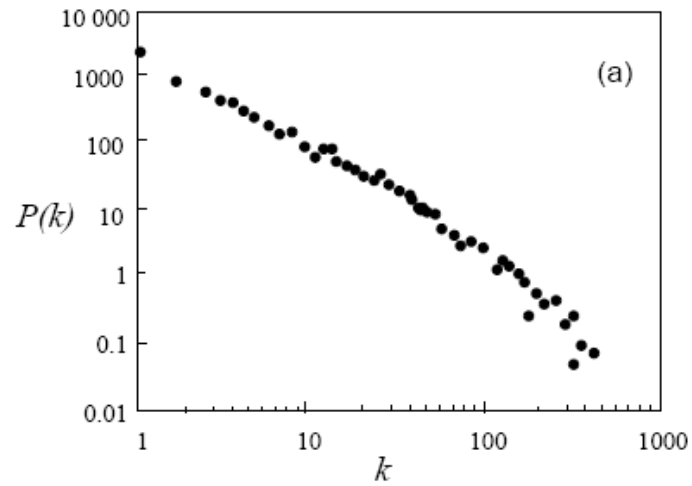


# E-mail hálózatok (Ebel 2002)

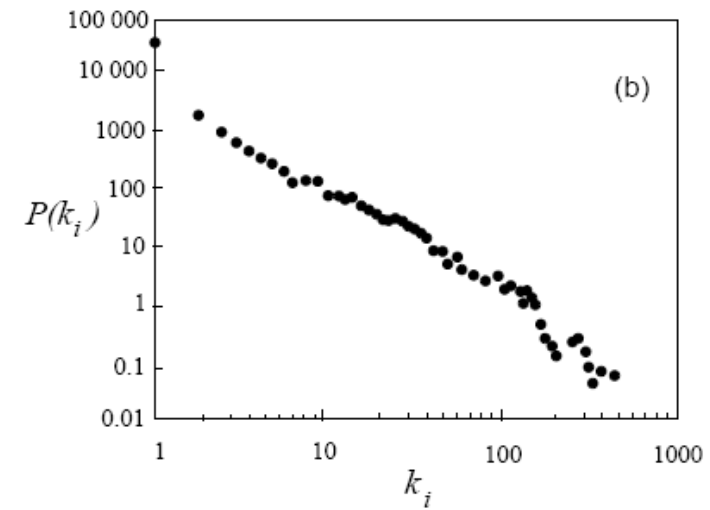
- Csomópontok: postafiókok
- Kapcsolatok: levelek
- Adatbázis: Kieleti egyetem hallgatóinak postafiókjairól készült logok
- 59.912 csomópont, 5165 belső
- $\langle k \rangle$  teljes 2.88  $\langle k \rangle$  belső  $\langle 25.45 \rangle$
- $C=0.156$  (irányítottság mellőzésére)
- $\langle l \rangle=4.95$



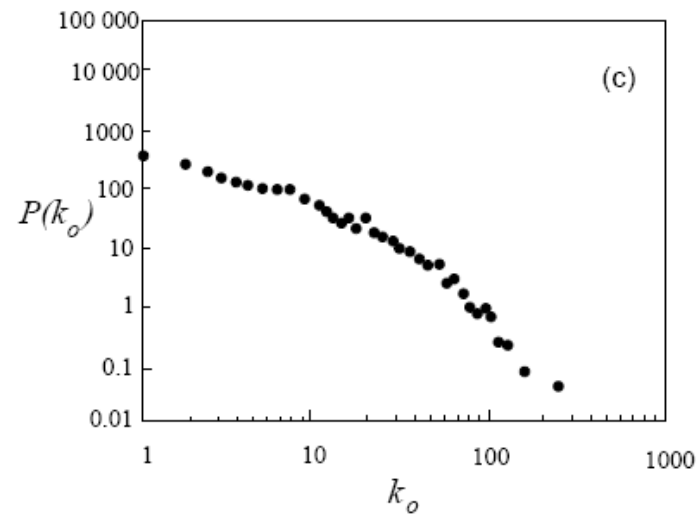
# E-mail hálózatok



$$\gamma = 1.3, k < 10^2$$



$$\gamma_i = 1.5, k < 10^2$$

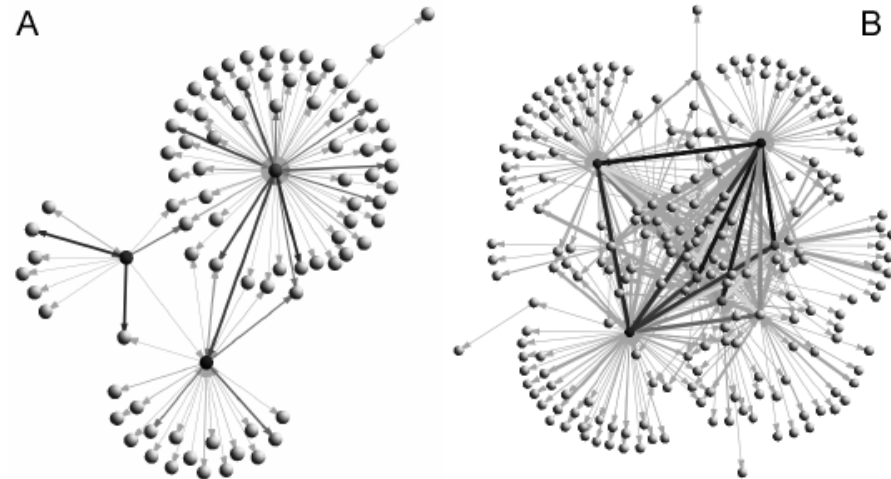


???

# Open source hálózatok (Valverde 2007)

- Csomópontok: programozók
- Kapcsolatok: e-mailek (kommunikáció adott idő alatt)
- Hálózat célja: nagy méretű szoftverek fejlesztése bottom-up
- Nincs központ
- Elosztott működés
  - Hibajavítás
  - Változtatások elfogadása, elvetése
- Amavis (email viruskereső) és TCL (scriptnyelv) közösségek
- Vegyes hálózat: önszerveződés és irányítás
- Magas reciprocitás

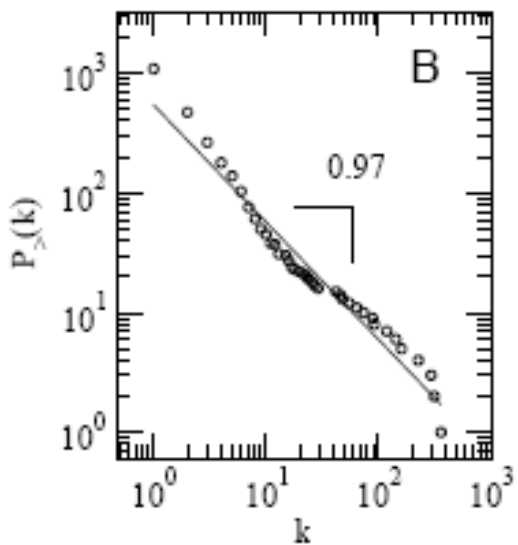
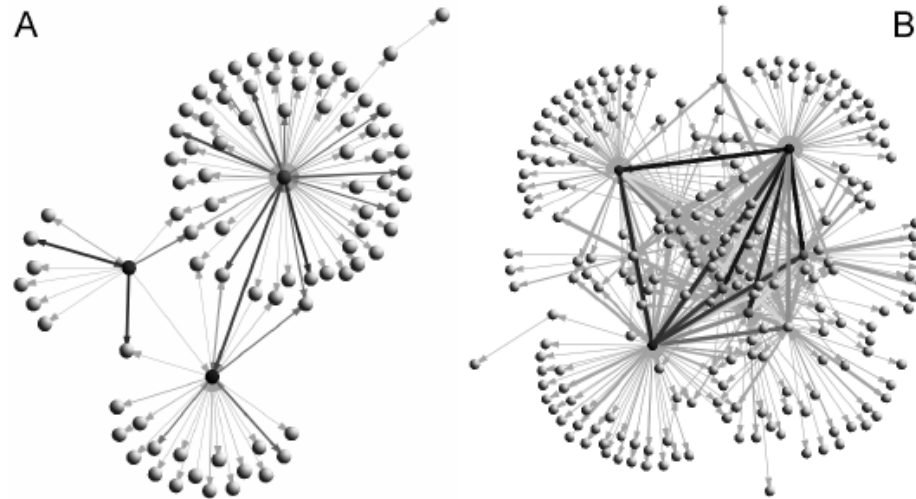
$$e_{ij} = \sum_{t=0}^T E_{ij}(t)$$





# Open source hálózatok

Project	$N$	$L$	$\rho^w$	$\langle k \rangle$	$\gamma$
Python	1090	3207	0.98	2.94	1.97
Gaim	1415	2692	0.98	1.9	1.97
Slashcode	643	1093	0.98	1.69	1.88
PCGEN	579	1654	0.98	2.85	2.04
TCL	215	590	0.98	2.74	1.97



$$\gamma \approx 2$$

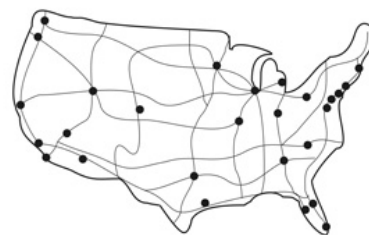
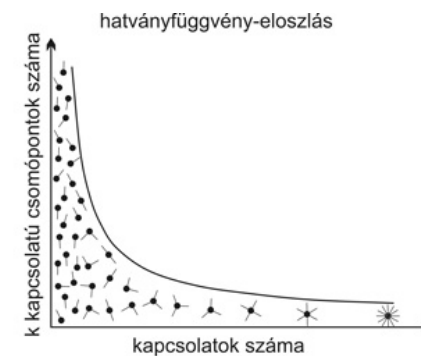
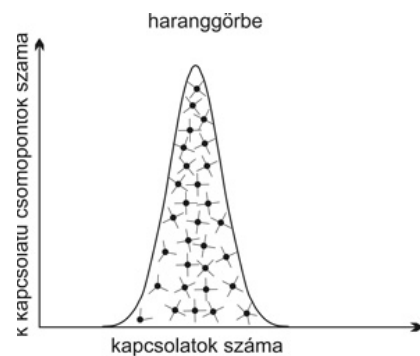
# Technológiai hálózatok

- Csomópontok: technikai berendezések
- Kapcsolatok: általában fizikai kapcsolat
- Cél: valamilyen szolgáltatás, vagy fizikai dolgok eljuttatása a felhasználókhoz
  - Áram
  - IP csomag
  - Maga a felhasználó (úthálózat, légitársaság, vasút stb.)

- Vasút:

N	$\langle k_i \rangle$	$\langle l \rangle$	C
587	66.79	2.16	0.7

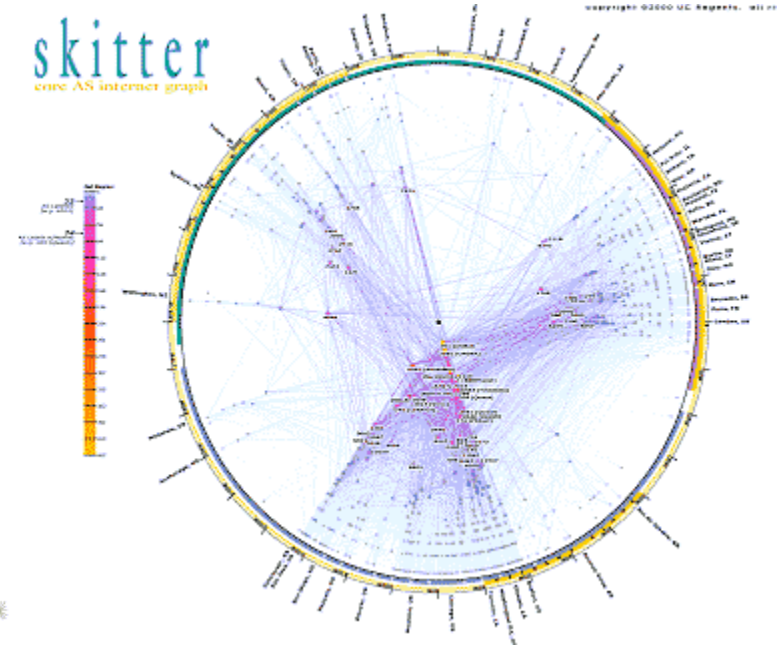
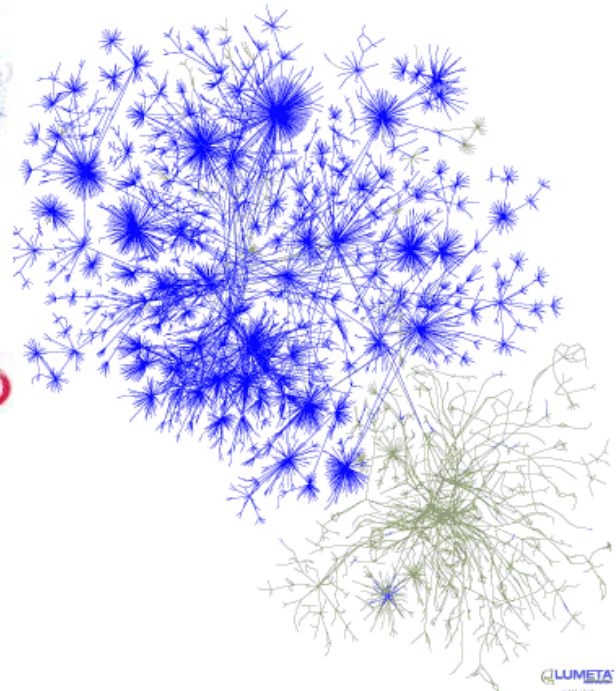
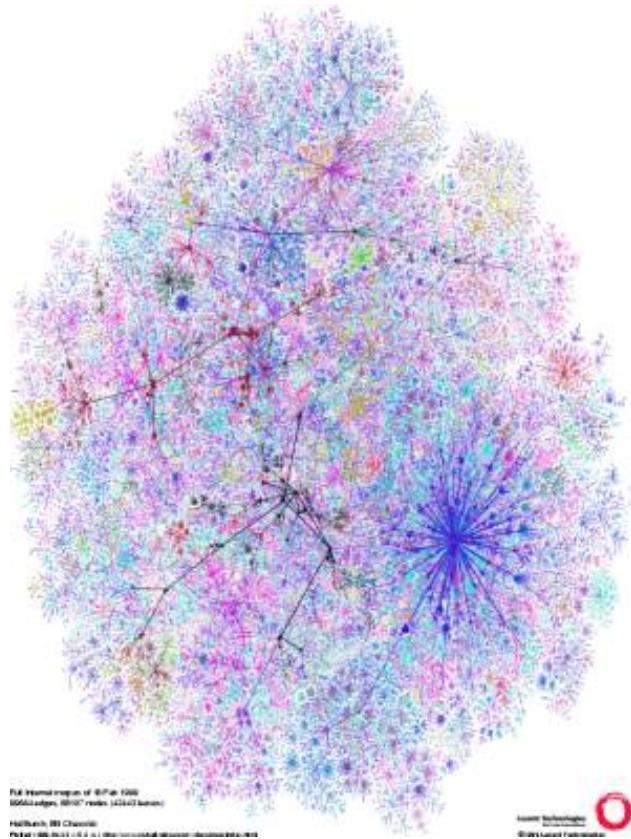
- Gyorsan lecsengő



# Villamos hálózat (Watts 1998)

- Csomópontok: transzformátorok, elosztóközpontok, generátorok
- Kapcsolatok: nagyfeszültségű távvezetékek
- Cél: áram elosztása, szállítása
- Adatbázis: Nyugati államok villamoshálózata
- $\langle k \rangle = 2.67 \rightarrow$  átlagos távolság viszonylag nagy
- Kisvilág?
- Nagyobb hálózat kellene, de nincs

Foksz	$N$	$L$	$C$	$C/C_r$	$\bar{l}$	$\bar{l}/\bar{l}_r$
Western States Power Grid	4941	$6.61 \times 10^3$	0.08	$1.5 \times 10^2$	18.7	1.5



# Internet topológia (domain szinten) (Faloutsos 1999)

- Csomópontok: autonóm rendszerek, routing szinten egységesen kezelt IP prefixek, melyet egy vagy több operátor üzemeltet
- Kapcsolatok: összeköttetések
- Cél: IP csomagok szállítása
- Adatbázis: NLANR (National Laboratory for Applied Network Research [www.nlanr.net](http://www.nlanr.net) 1997 óta)

• Átlagos

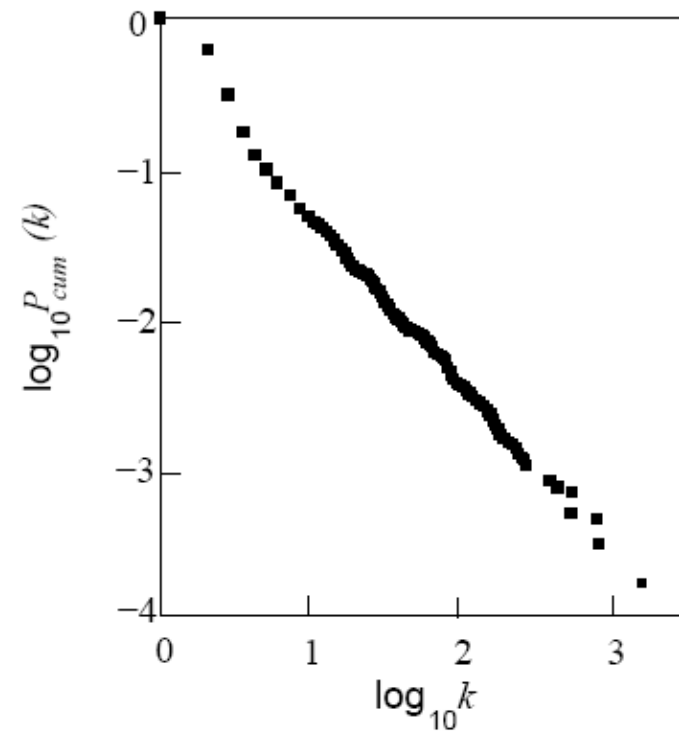
• Növeks

	$N$	$L$	$\bar{k}$	$\bar{\ell}$	$C$	$N_{new}$	$N_{died}$
November 1997	3015	5156	3.42	3.76	—	—	—
Average 1997	3112	5450	3.5	3.8	0.18	309	129
April 1998	3530	6432	3.65	3.77	—	—	—
Average 1998	3834	6990	3.6	3.8	0.21	1990	887
December 1998	4389	8256	3.76	3.75	—	—	—
Average 1999	5287	10 100	3.8	3.7	0.24	3410	1713



# Internet domain szinten

- Dinamizmus
- Gazdag még gazdagabb

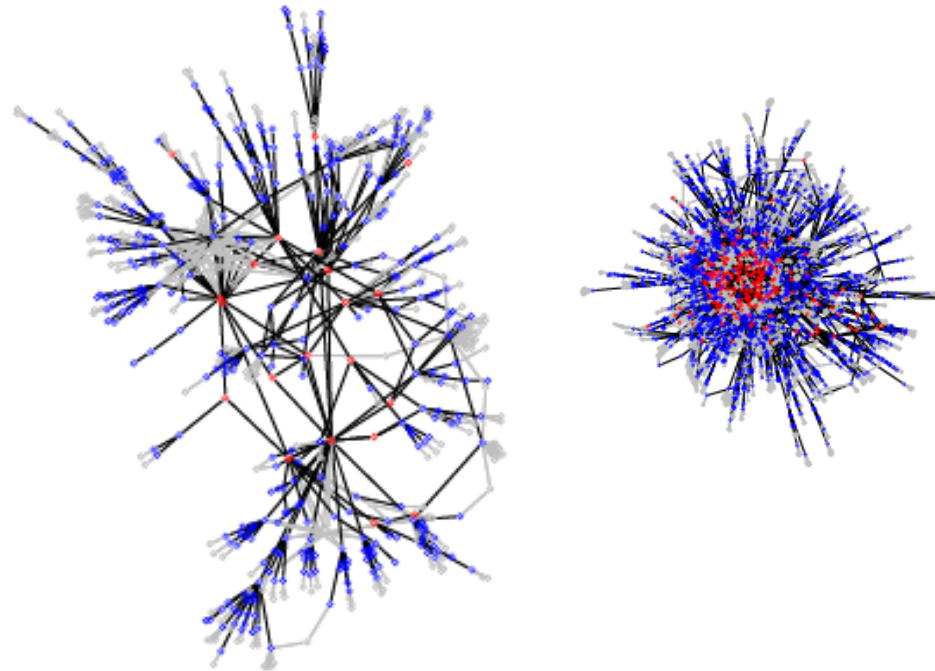
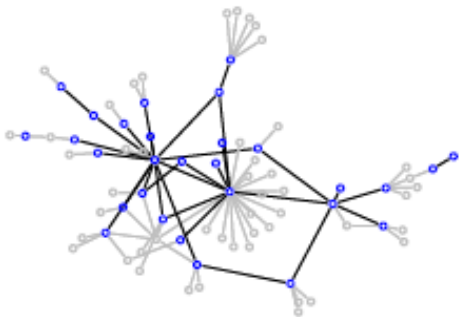


$$\gamma = 2.2$$

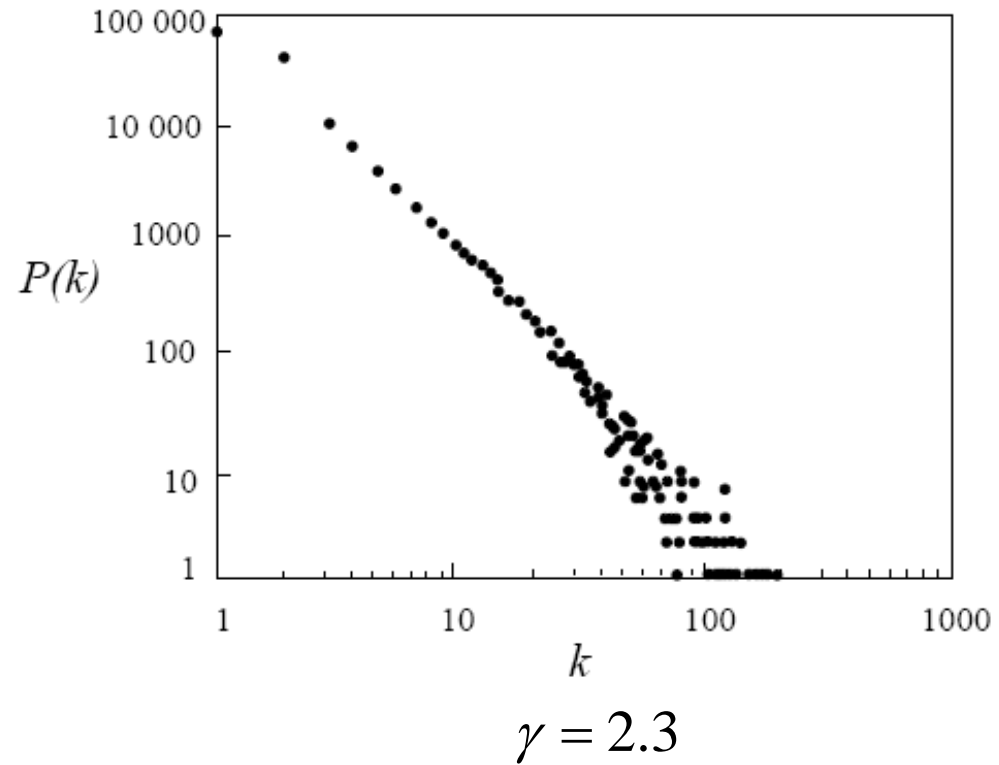
# Internet router szinten (Govindan 2000)

- Csomópontok: routerek
- Kapcsolatok: 1 IP hop
- Sajátosságok: aggregáló csomópontok (nem igazi routerek), tényleges útválasztók
- Topológia előállítása nem triviális
- Traceroute DNS-ből vett címekhez

N	$\langle k \rangle$	$\langle l \rangle$
150.000	2.66	11



# Internet router szinten



- Gazdag még gazdagabb?

# World wide web (Barabási 1999)

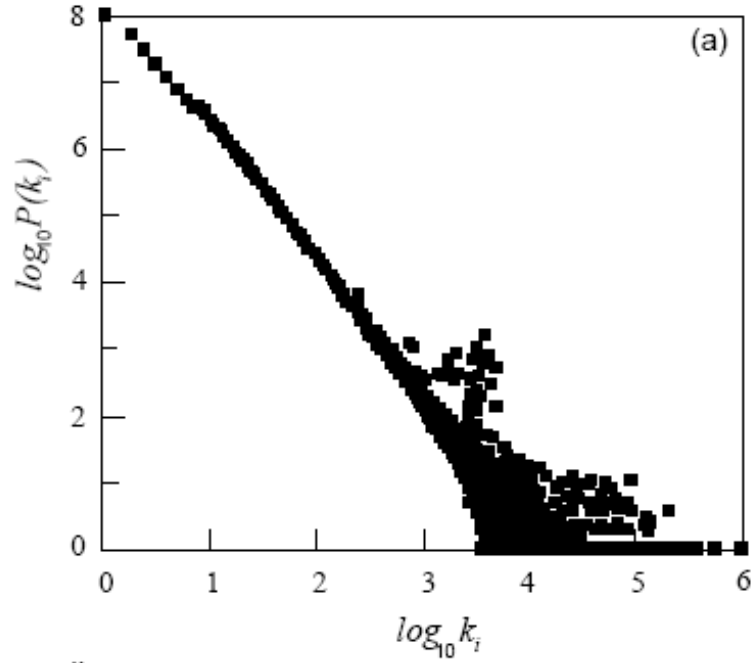
- Csomópontok: weboldalak
- Kapcsolatok: linkek
- Irányított gráf
- Adatbázis nd.edu körzet
- Többszöri mérések az időszakban

N	$\langle k \rangle$	$\langle l \rangle$	C
325.729	7.85	11.2	0.29

	N	E
1999 május	203 millió	1466 millió
1999 október	271 millió	2130 millió

- Nemlineáris növekedés

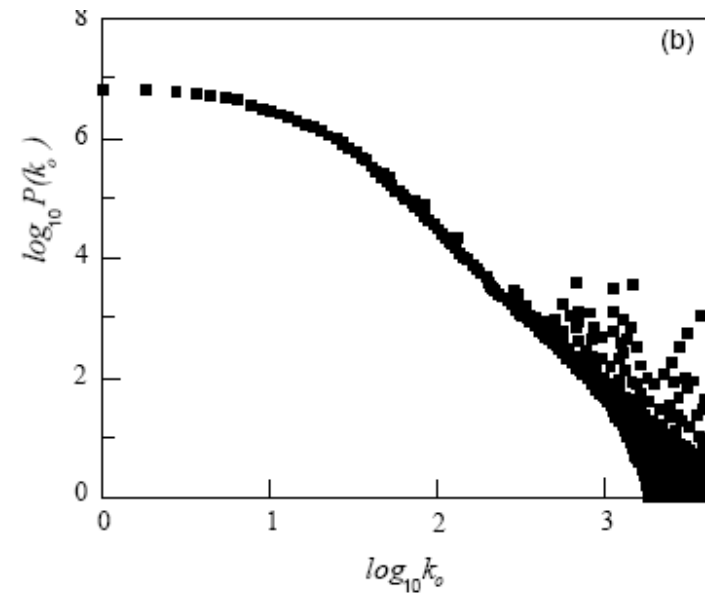
# WWW



$$\gamma_i = 2.1$$

$$\gamma_o = 2.7$$

- Többféle illesztés van



# WWW átmérője (Barabási 1999)

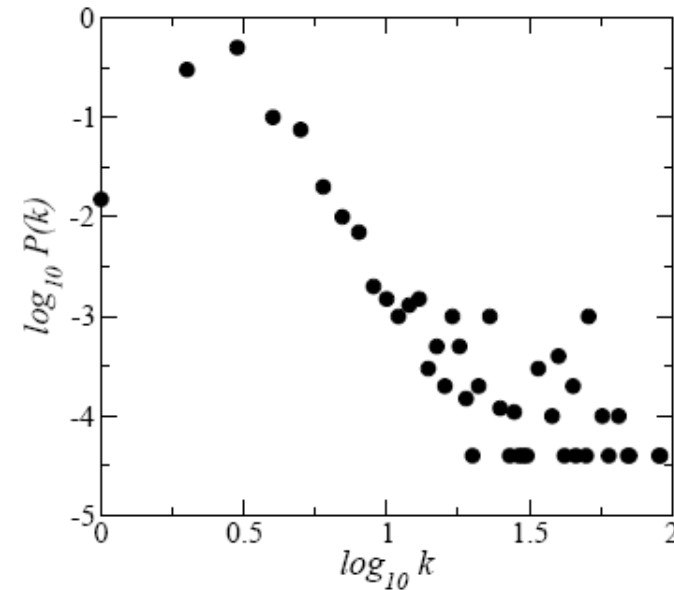
- Eljárás kidolgozása a teljes web átmérőjének becslésére
  - Fraktálok
  - Fokszámeloszlás mérése az nd.edu tartományban
  - Hálózatgenerálás ezzel  $\bar{\ell}_d(N) \approx 0.35 + 2.06 \log_{10} N$
  - Trend meghatározása
- 
- Teljes webre az átmérő 19 (1999-ben kb. 800.000.000 oldal)



# Elektornikus áramkörök (Solé 2002)

- Csomópontok: áramköri elemek (tranzisztor, dióda, kondenzátor, logikai kapuk stb.)
- Kapcsolatok: huzalok
- Régebbi áramkörök néhány száz elem
- Ma már lehetséges statisztikát is csinálni
- Gazdag még gazdagabb?

N	$\langle k \rangle$	$\langle l \rangle$	C
20.000	4	11.05	0.03

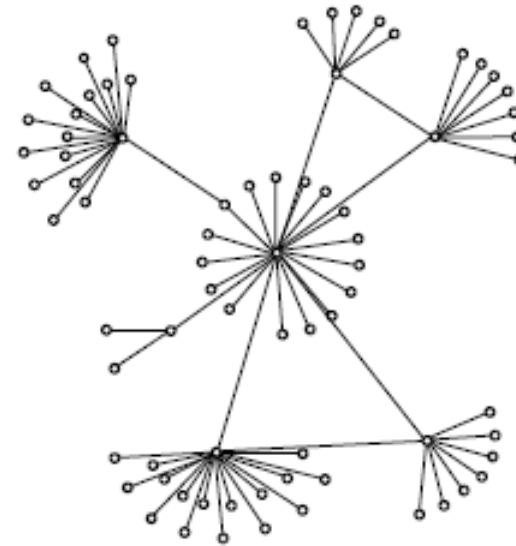


$$\gamma \approx 3?$$

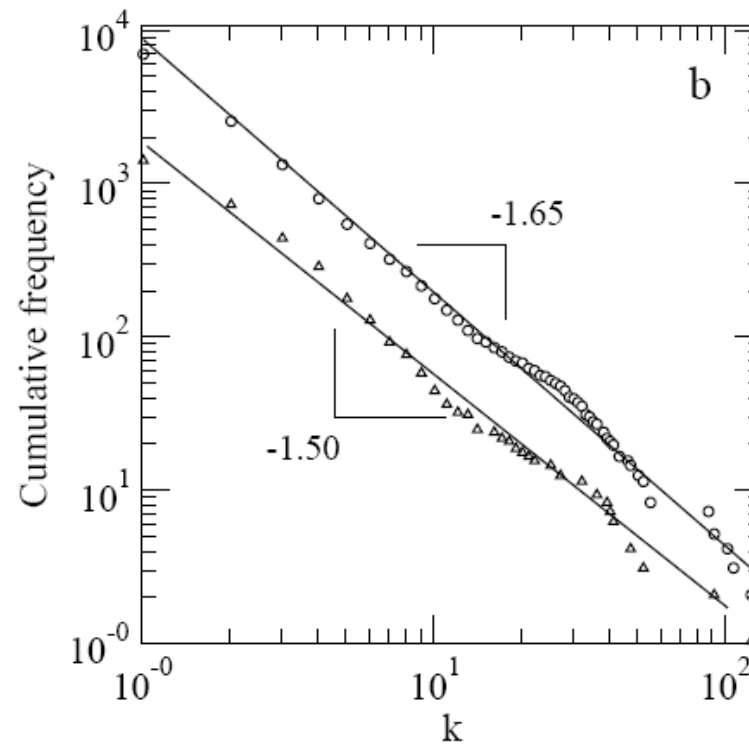
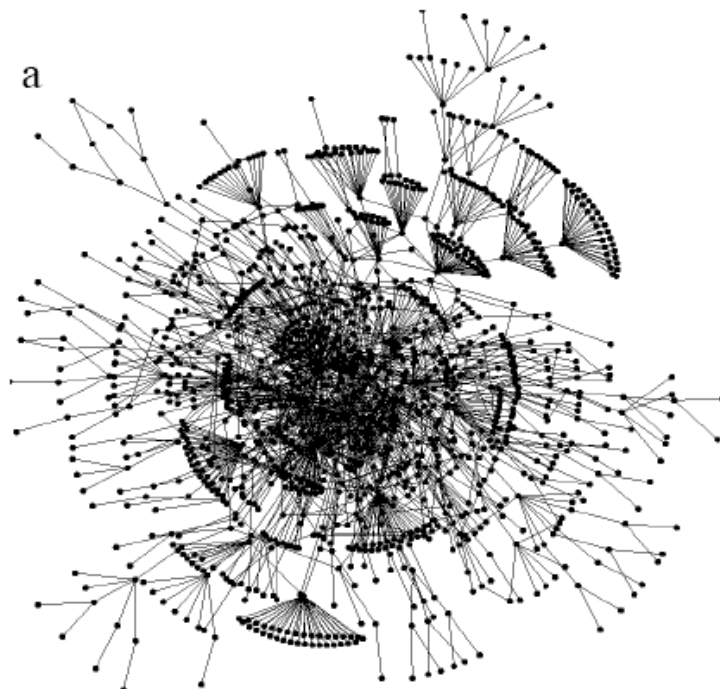
# Számítógépes játék, java development framework (Sole 2002)

- Szoftverek egyre komplexebbek → sok modul, bonyolult funkciók → rengeteg bug
- Moduláris struktúra → osztályok → osztálydiagramok (Objektumorientált paradigma)
- Csomópontok: modulok
- Kapcsolatok: modulok közötti interakciók (üzenetváltás, függvényhívás stb.)
- Java: 9257 modul, ebből 3115 összefüggő komponens a legnagyobb 1376 modul 2174 él
- Nemlineáris növekedés

N	$\langle l \rangle$	C
1376	6.39	0.06



# JDK



Tervezett hálózat (optimalizálás)

Önszerveződés?

Gazdag még gazdagabb?

$$\gamma = 2.5$$

# Játék osztálydiagramja

- Grafika
- Szimuláció
- Hangok és zene
- Memória menedzsment

N	E	<l>	C
1989	4780	6,2	0.08

$$\gamma \approx 2.85$$

- Első cikk arról, hogy a skálafüggetlen szerkezet optimalizáció közben is megjelenhet
- Akkor most önszerveződés vagy nem?

# Technológiai hálózatok

- Skálafüggetlen szerkezet mögötti törvényszerűségek
  - Természettudományos magyarázat: önszerveződés, megjelenő tulajdonságok
  - Mérnöki magyarázat: optimalizáció
- Nem eldöntött kérdések

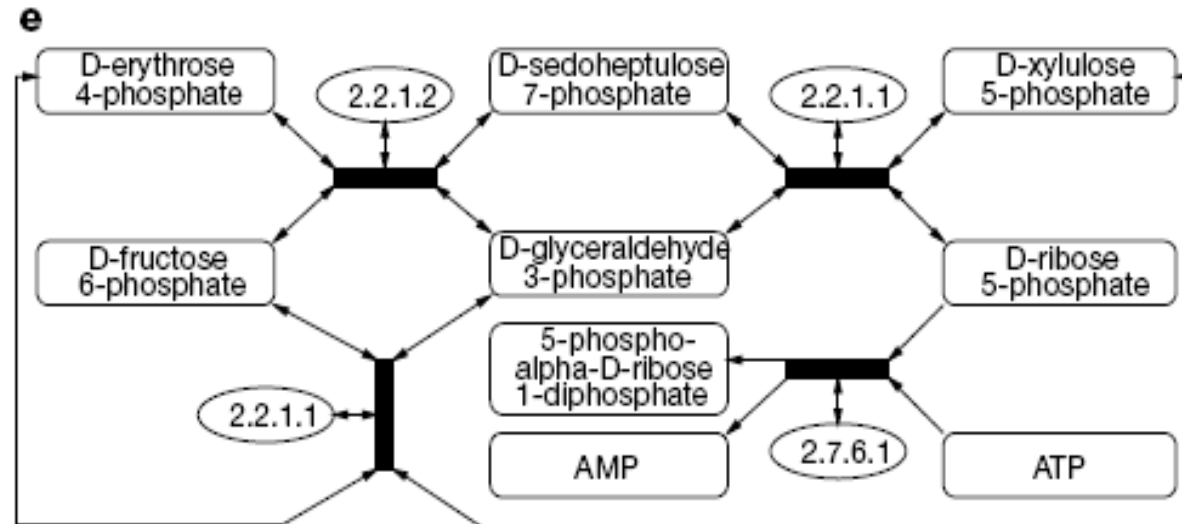
# Biológiai hálózatok

- Csomópontok: élő organizmusok, vagy molekulák
- Kapcsolatok: kémiai reakciók, fizikai kapcsolódás, szabályozás stb.
- Anyagcsere hálózatok
- Sejthálózat, teljes molekuláris felépítés (gének, fehérjék, egyéb molekulák)
- 2000 jún. 26. emberi genom feltérképezése (hárommilliárd AGCT)
  - értelmezni komoly kihívás

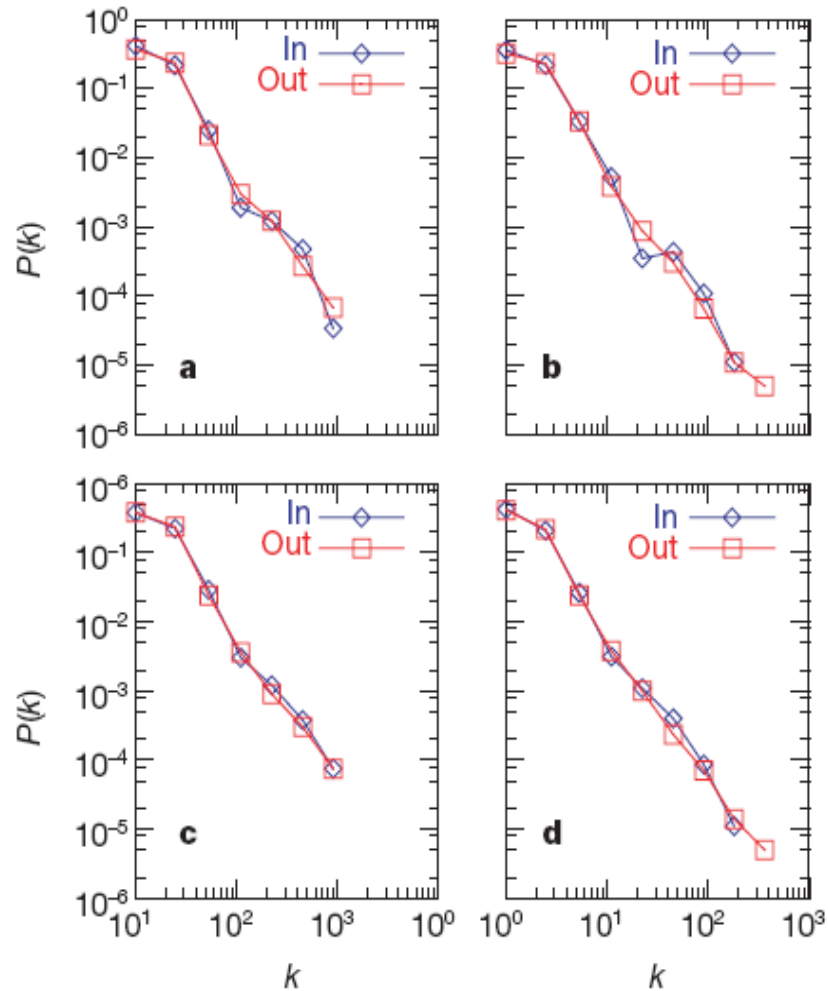


# Anyagcsere hálózatok (Barabási, Oltvai 2000)

- Csomópontok: molekulák
- Kapcsolatok: biokémiai reakciók
- 43 organizmus esetében vizsgálták
- E.coli baktérium

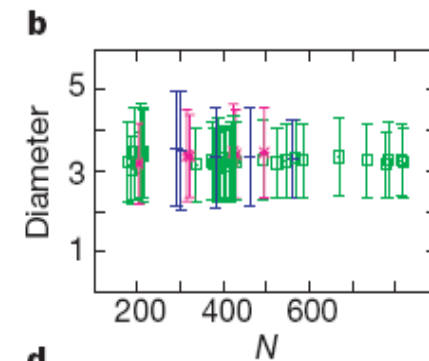
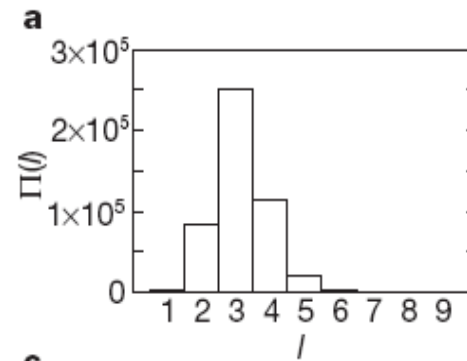


# Anyagcsere hálózat



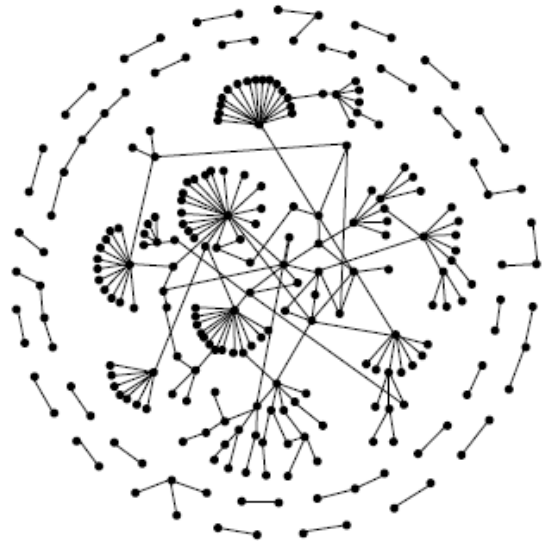
$$\gamma_{in} = 2.2$$

$$\gamma_{out} = 2.2$$



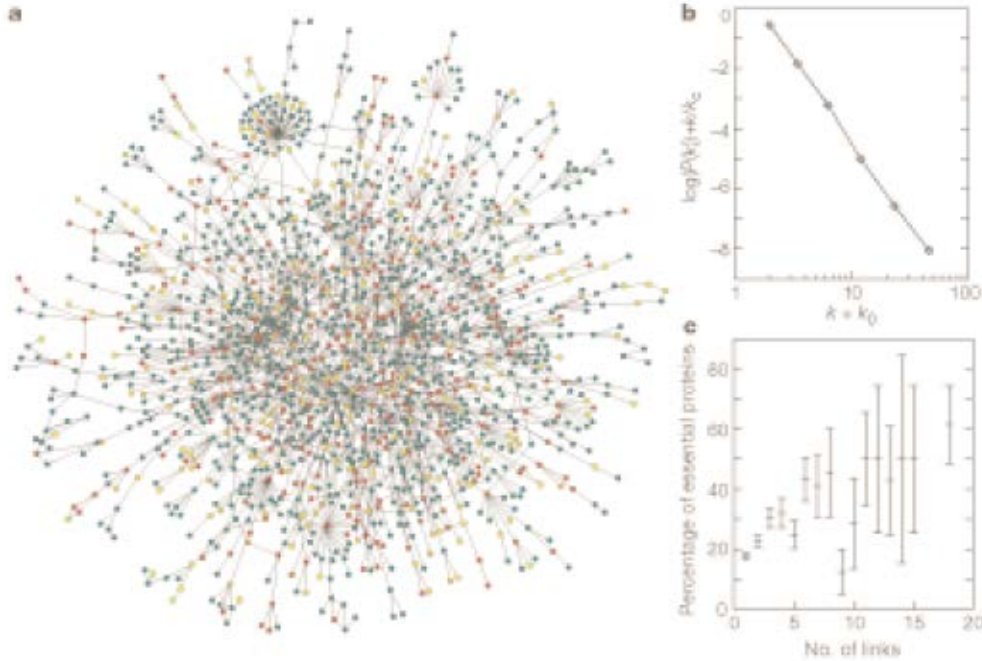
# Fehérjehálózatok (Barabási, Oltvai 2001)

- Kétféle van
- 1: Fehérjék fizikai kapcsolati hálója
- Csomópontok: fehérjék
- Kapcsolatok: fizikai összeköttetések (hemoglobin)
- 2: Fehérjék szabályozási hálója
- Csomópontok: fehérjék
- Kapcsolatok: szerkezet



tott)

# Fehérjehálózatok



$$\gamma = 2.5$$

N	E	$\langle l \rangle$
1870	2240	6,8

*S. Cerevisiae* (sörélesztő)

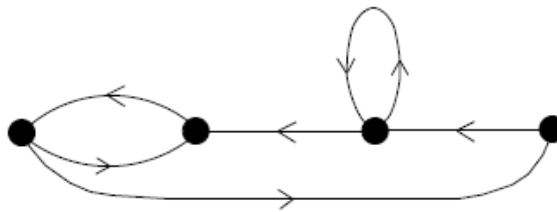
Gazdag még gazdagabb?

# Gének-fehérjék szabályozási háló

- Elképesztően bonyolult, most is folyik a feltérképezés
- Élesztő 6300 gén
- C. elegans 20.000 (300 neuron)
- Ember 30.000 (korábbi becslés legalább 100.000) (1 milliárd neuron)
- Élet könyve után, most jön az Élet térképe
- Genetikai betegségek kezelése
- Hatékonyabb gyógyszerek?

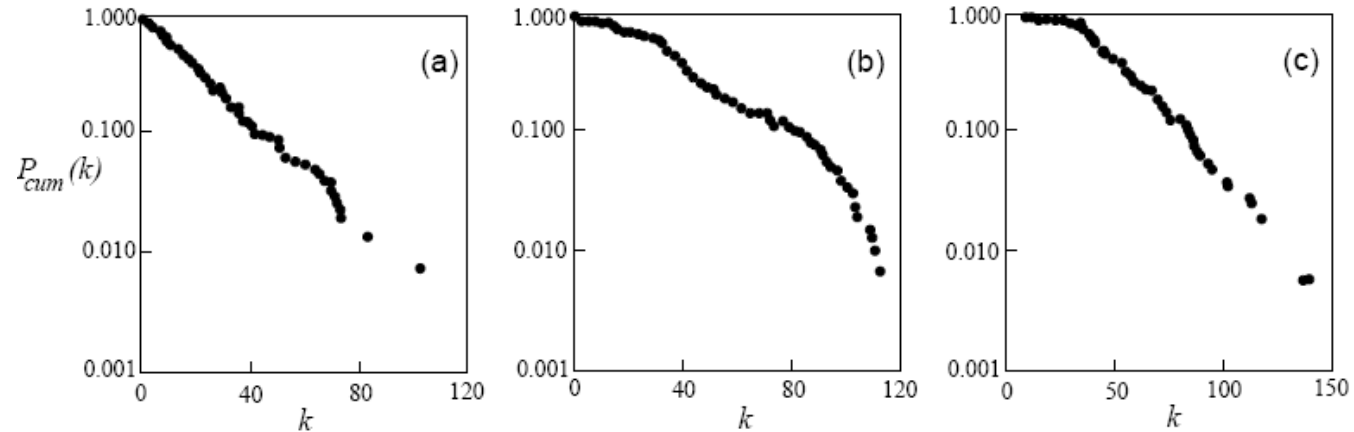
# Ökológiai hálózatok (Dunne, Williams 2002)

- Táplálkozási láncok
- Csomópontok: élőlények
- Kapcsolatok: ki esz meg kit
- Ökoszisztéma komplex biológiai egyensúlya
- Nagyon nehéz előállítani
- Egyelőre kis méretűek vannak





# Ökológiai hálózatok



Food web	$N$	$L$	$L/N^2$	$C$	$\bar{\ell}$
El Verde Rainforest	155	$1.51 \times 10^3$	0.026	0.12	2.20
Lake Tahoe	172	$3.88 \times 10^3$	0.131	0.14	1.81
Mirror Lake	172	$4.32 \times 10^3$	0.146	0.14	1.76

tengeri vidra → tengeri sün → moszatok → halak és erózió

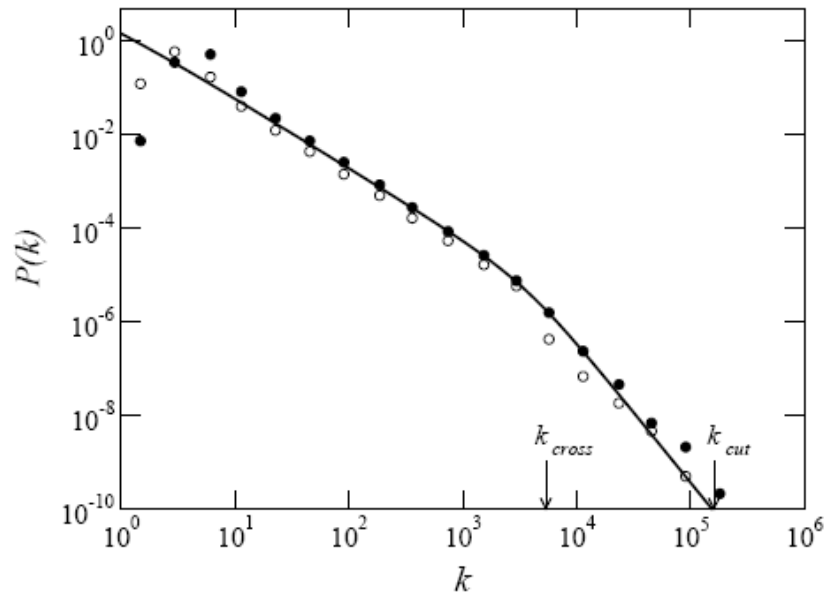
kecske és a parlagfű, vagy a méhek

# Word web (Solé 2001)

- Szövegek statisztikus elemzésének nagy múltja van
- Szavak előfordulásának gyakorisága
- Szavak előfordulási hálózata
- Csomópontok: szavak
- Kapcsolatok: közös előfordulás gyakorisága
- Angol szövegeket vizsgált
- 470.000 szó
- Kapcsolat, ha egyszer is interakcióban volt a két szó
- Interakció: egymás mellett, vagy egymástól két szó távolságra
  - Többféle szabály lehet, az eredményeket nem nagyon befolyásolja

N	<k>	<l>	C
470.000	72	2.65	0.69

# Word web



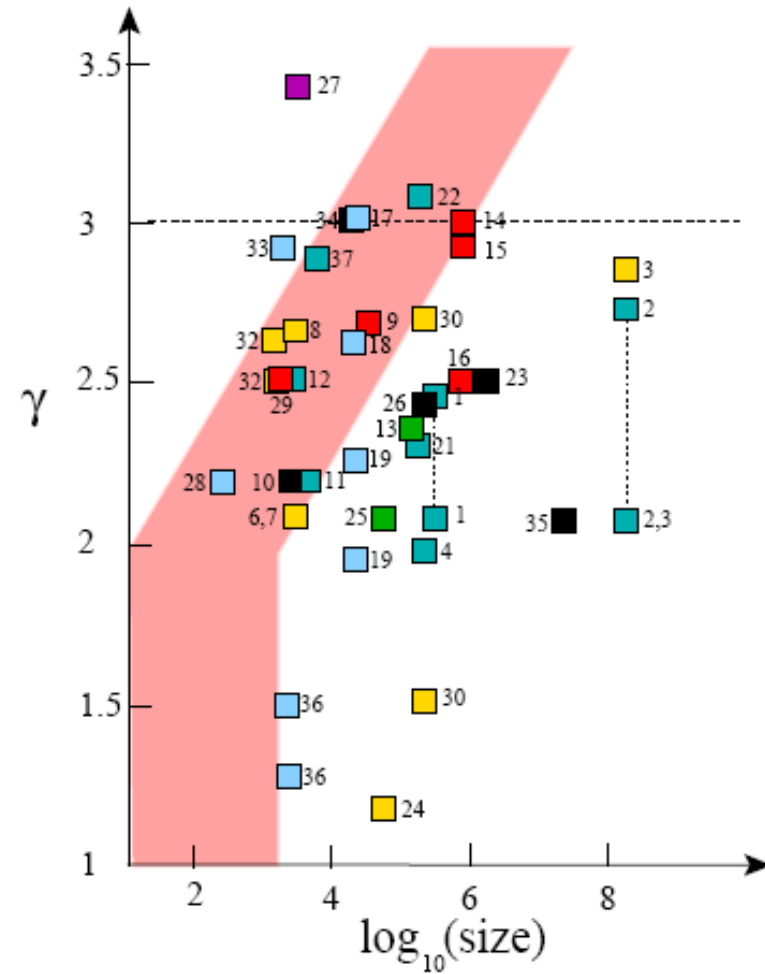
$$\gamma_1 = 1.5$$

$$\gamma_2 = 2.7$$

# Áttekintés

Network or subgraph	Number of vertices	Number of edges	$\gamma$	$C$	$C/C_r$	$\bar{\ell}$	$\bar{\ell}/\bar{\ell}_r$	References	
1 Complete map of the nd.edu domain of the Web	325 729	1 469 680	$\gamma_i = 2.1$ $\gamma_o = 2.4\bar{4}$	1	—	—	11.2	—	Albert <i>et al.</i> 1999
2 Pages of the WWW scanned by Altavista <sup>1</sup> in October 1999	$2.711 \times 10^8$	$2.130 \times 10^9$	$\gamma_i = 2.1$ $\gamma_o = 2.7$	2	—	—	16 (6.8)	(1.0)	Kumar <i>et al.</i> 2000a, Broder <i>et al.</i> 2000
3 '————' (another fitting of the same data)	—	—	$\gamma_i = 2.10$ $\gamma_o = 2.8\bar{2}$	3	—	—	—	—	Newman <i>et al.</i> 2001
4 Map of sites of a domain of the WWW, spring 1997	$2.60 \times 10^5$	—	$\gamma_i = 1.94$	4	—	—	—	—	Adamic <i>et al.</i> 2000
5 Undirected map of sites in a domain of the WWW	153 127	$2.70 \times 10^6$	—	5	0.108	$0.47 \times 10^3$	3.1	0.93	Adamic 1999
6 A set of public company home pages	4923	$1.335 \times 10^7$	$\gamma_i = 2.05$	6	—	—	—	—	Pennock <i>et al.</i> 2002
7 A set of US newspaper home pages	—	—	$\gamma_i = 2.05$	7	—	—	—	—	Pennock <i>et al.</i> 2002
8 A set of university home pages	—	—	$\gamma_i = 2.63$	8	—	—	—	—	Pennock <i>et al.</i> 2002
9 A set of computer scientist home pages	56 880	—	$\gamma_i = 2.66$	9	—	—	—	—	Pennock <i>et al.</i> 2002
10 Interdomain level of the Internet, December 1998	4389	8256	2.2	10	—	—	4	0.6	Faloutsos <i>et al.</i> 1999
11 Interdomain level of the Internet, December 1999 <sup>2</sup>	6374	13 641	2.2	11	0.24	$3.3 \times 10^2$	3.7	0.58	Pastor-Satorras <i>et al.</i> 2001
12 Router level of the Internet in 1995	3888	5012	2.5	12	—	—	12.1	1.39	Faloutsos <i>et al.</i> 1999
13 Router level of the Internet in 2000	$\sim 150\,000$	$\sim 200\,000$	$\sim 2.3$	13	—	—	10	0.8	Govindan <i>et al.</i> 2000
14 Citations in the ISI database 1981–June 1997	783 339	6 716 198	$\gamma_i = 3.0$	14	—	—	—	—	Redner 1998
15 '————' (another fitting of the same data)	—	—	$\gamma_i = 2.9$	15	—	—	—	—	Tsallis <i>et al.</i> 2000
16 '————' (another estimate from the same data)	—	—	$\gamma_i = 2.5$	16	—	—	—	—	Krapivsky <i>et al.</i> 2000
17 Citations in <i>Phys. Rev. D</i> 11–50 (1975–1994)	24 296	351 872	$\gamma_i = 3.0$	17	—	—	—	—	Redner 1998
18 '————' (another fitting of the same data)	—	—	$\gamma_i = 2.6$	18	—	—	—	—	Tsallis <i>et al.</i> 2000
19 '————' (another estimate from the same data)	—	—	$\gamma_i = 2.3$	19	—	—	—	—	Krapivsky <i>et al.</i> 2000
20 '————' (another estimate from the same data)	—	—	$\gamma_i = 1.9$	20	—	—	—	—	Vázquez 2001b
21 Collaboration network of movie actors	212 250	61 085 555	2.3	21	—	—	4.54	1.25	Barabási <i>et al.</i> 1999
22 '————' (another fitting of the same data)	—	—	3.1	22	—	—	—	—	Albert <i>et al.</i> 2000a
23 Collaboration network of MEDLINE	1 388 989	$1.028 \times 10^7$	2.5	23	0.066	$6 \times 10^3$	4.6	0.9	Newman 2001e
24 Coauthorships in the SPIRES <sup>3</sup> e-archive	56 627	$4.898 \times 10^6$	1.2	24	0.726	$0.24 \times 10^3$	4.0	1.88	Newman 2001e
25 Collaboration net collected from math. journals	70 975	$0.132 \times 10^6$	2.1	25	0.59	$1.1 \times 10^4$	9.5	1.16	Barabási <i>et al.</i> 2002
26 Collaboration net collected from neurosci. journals	209 293	$1.214 \times 10^6$	2.4	26	0.76	$1.4 \times 10^4$	6	1.2	Barabási <i>et al.</i> 2002
27 Web of human sexual contacts <sup>4</sup>	2810	—	3.4	27	—	—	—	—	Liljeros <i>et al.</i> 2001
28 Networks of metabolic reactions <sup>5</sup>	$\sim 200\text{--}800$	$\sim 600\text{--}3000$	$\gamma_i = 2.2$ $\gamma_o = 2.2$	28	0.32	12	3.2	0.95	Jeong <i>et al.</i> 2000, Wagner <i>et al.</i> 2001
29 Net of protein–protein interactions (yeast proteome) <sup>6</sup>	1870	2240	$\sim 2.5$	29	0.022	4.4	6.8	0.8	Jeong <i>et al.</i> 2001, Wagner 2001a
30 Word Web <sup>7</sup>	470 000	17 000 000	1.5 / 2.7	30	0.69/0.44	$4.4 \times 10^3/2.8 \times 10^3$	2.65	0.87	Ferrer i Cancho <i>et al.</i> 2001a
31 Food web of Silwood park <sup>8</sup>	154	366	$\sim 1$	31	0.15	5	3.4	1.05	Montoya <i>et al.</i> 2001, Solé <i>et al.</i> 2002
32 Java Development Framework (largest component)	1376	2174	2.5	32	0.06	25	6.39	1.02	Valverde <i>et al.</i> 2002
'————' (second largest component)	1364	1930	2.65	—	0.08	40	6.91	1.01	Valverde <i>et al.</i> 2002
33 Computer game graph	1989	$4.78 \times 10^3$	2.85	33	0.08	35	6.2	1.28	Valverde <i>et al.</i> 2002
34 Large digital electronic circuits	$2 \times 10^4$	$4 \times 10^4$	3.0	34	$3 \times 10^{-2}$	$1.5 \times 10^2$	$\sim 6$	$\sim 1$	Ferrer i Cancho <i>et al.</i> 2001b
35 Telephone call graph <sup>9</sup>	$47 \times 10^6$	$8 \times 10^7$	$\gamma_i = 2.1$	35	—	—	—	—	Aiello <i>et al.</i> 2000
36 E-mail net (student accounts in Kiel University) <sup>10</sup>	5165	$6.57 \times 10^4$	1.3, $\gamma_i=1.5$	36	0.156	$3.25 \times 10^3$	4.95	0.48	Ebel <i>et al.</i> 2002
37 Energy landscape network for a 14-atom cluster	4196	87 219	2.78	37	0.073	7.4	2.32	1.04	Doye 2002

# Áttekintés



# Szűkebb értelemben vett komplex hálózatok

- **Speciális értelemben**

1. Nem teljesen véletlenszerű kapcsolatok, "csoportosuló"
2. Kis átmérő, rövid utak, kisvilág
3. Skálafüggetlen szerkezet: preferencia alapú kapcsolódás, gazdag még gazdagabb



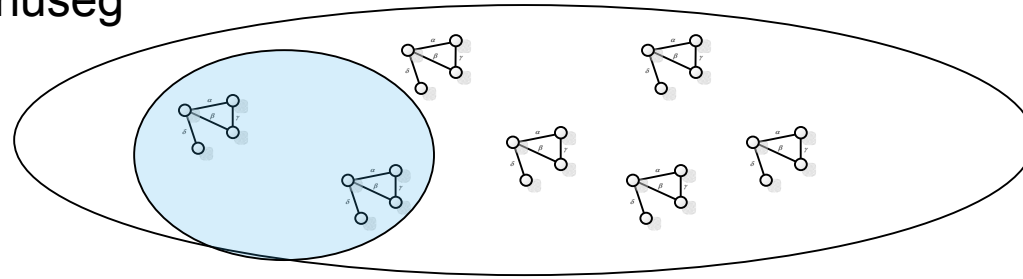


# Nagy hálózati modellek

# Véletleg gráfok definíciója

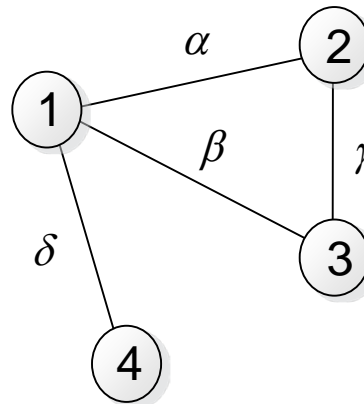
- 1. Definíció

- $G(N,p)$  véletlen tér ( $S\{\Omega, 2^\Omega, P(2^\Omega) \rightarrow \mathbb{R}\}$ ) /itt most diszkrét/
  - $N$ : csomópontok száma
  - $p$ : élvalószínűség



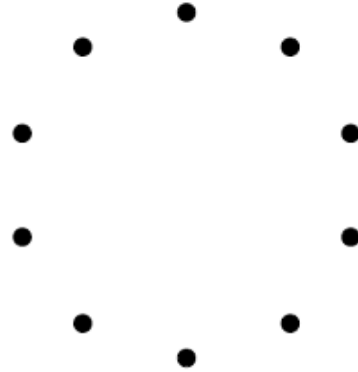
- 2. Definíció

- Olyan gráf, melyben a csomópontok közötti élek azonos  $p$  valószínűséggel vannak behúzva: élek szám  $p$ -től függő véletlen változó

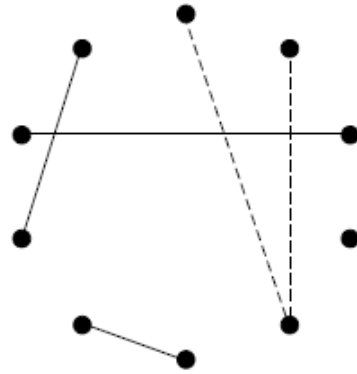


# Véletlen gráf példa N=10-re

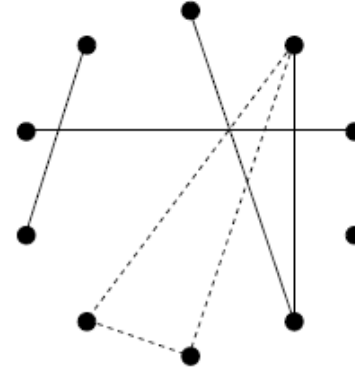
$p=0$



$p=0.1$



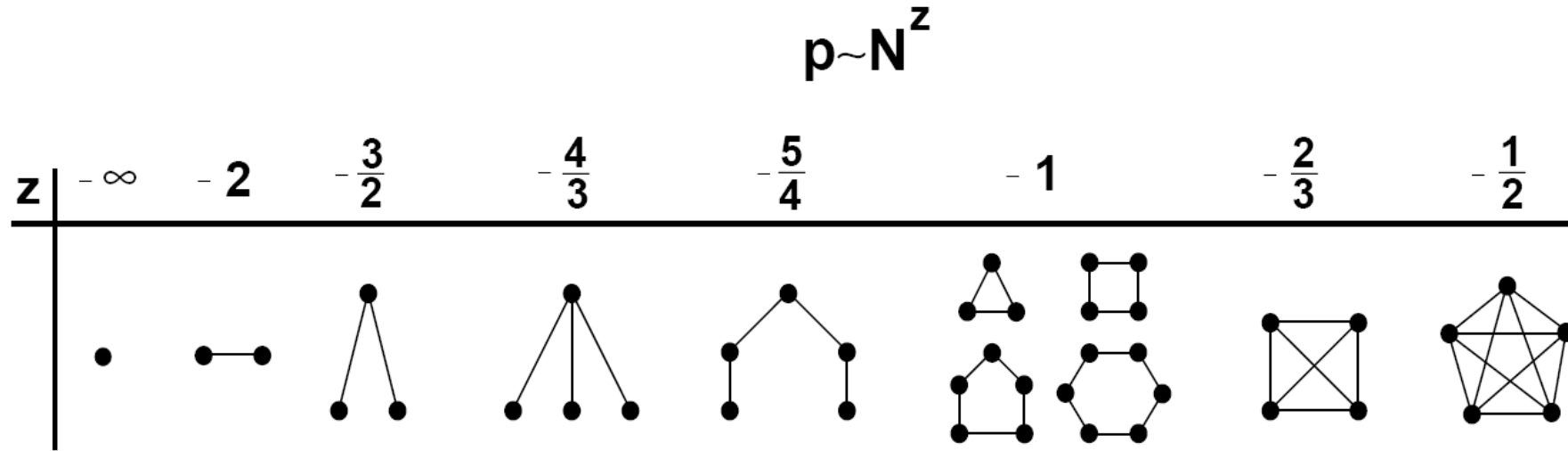
$p=0.15$



# Random hálózat fázisátalakulása



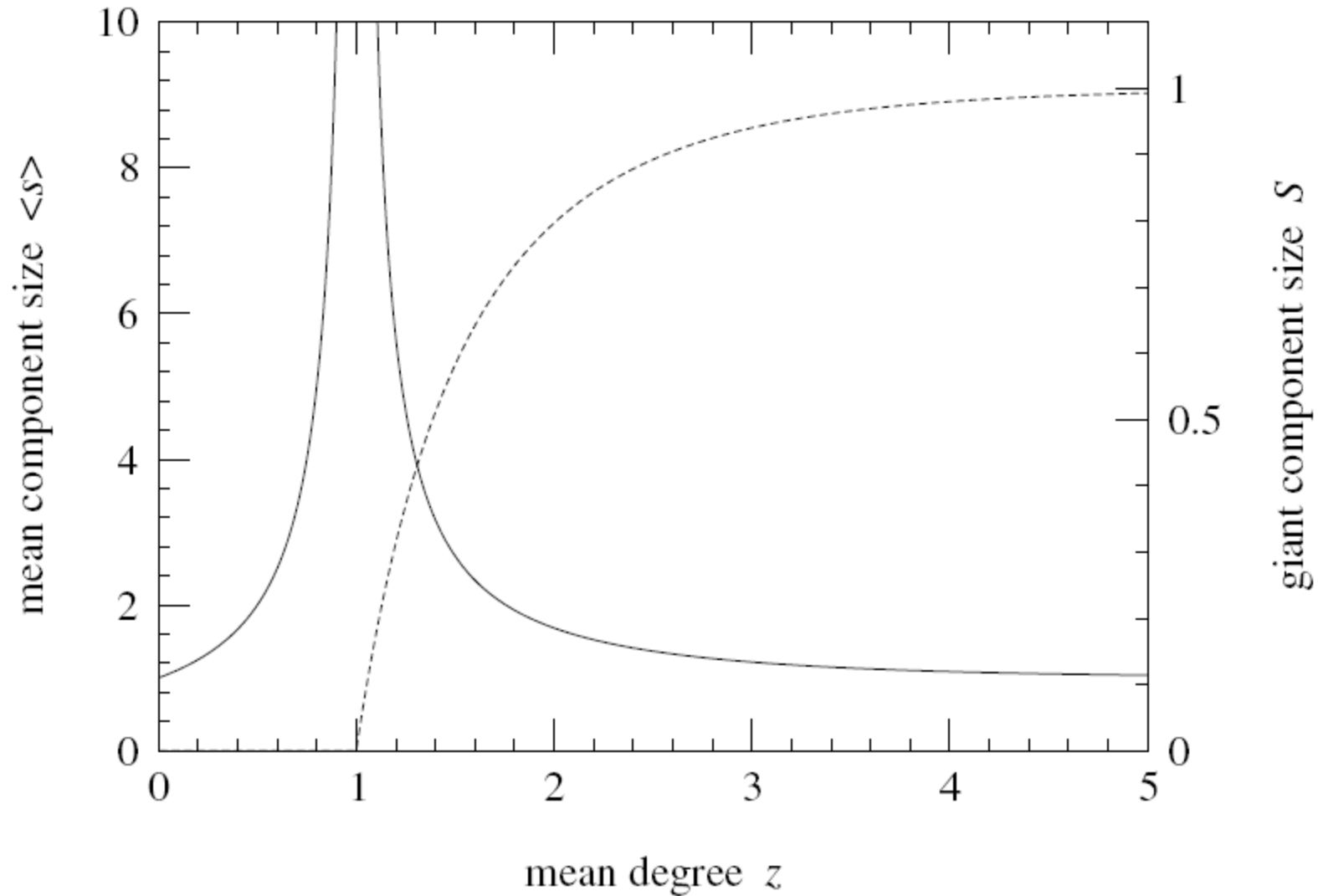
# Gráf evolúciója



• N=100 esetén:

z	$-\infty$	-2	$-3/2$	$-4/3$	$-5/4$	-1	$-2/3$	$-1/2$	0
p	0	0.0001	0.001	0.0021	0.00398	0.01	0.0464	0.1	1

# Az óriás komponens megjelenése: fázisátmenet



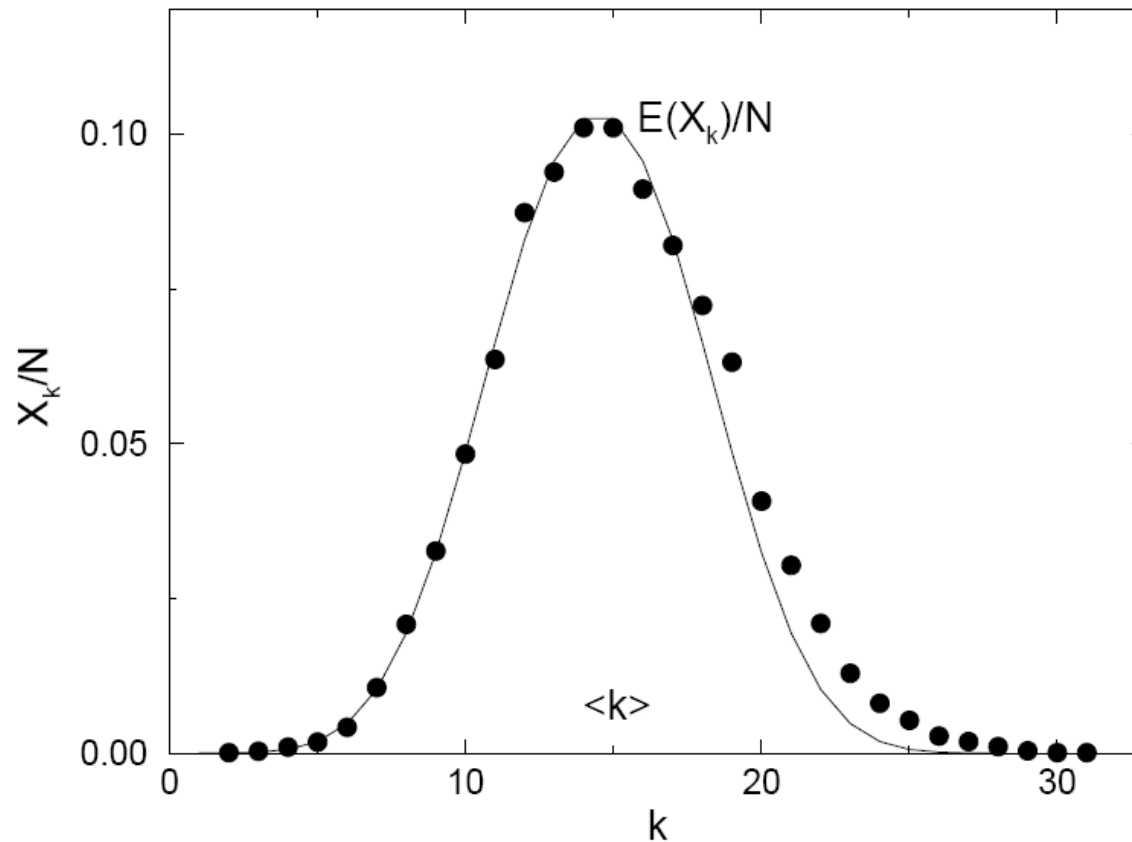


# Összefüggőség

- A gráf átlagos fokszáma  $\langle k \rangle = 2n/N = p(N-1) \sim pN$
- Amikor  $p \ll 1/N$  a különböző komponensek fák
- $\langle k \rangle \sim 1$  esetén a legnagyobb klaszter mérete  $N^{2/3}$ , az előző példa esetén: 21,54
- $\langle k \rangle \gg 1$  esetén a legnagyobb klaszter mérete  $[1 - f(\langle k \rangle)]N$
- az  $f(\langle k \rangle)$  exponenciális sebességgel 0 lesz...
- $p = 1/N$ -nél tehát megjelenik az óriás komponens
- $p = \ln(N)/N$  –nél a gráf összefüggővé válik

# Fokszámeloszlás példa: $N=10000$ , $p=0.0015$

$$P(k) = C_{N-1}^k p^k (1-p)^{N-1-k}$$
$$P(k) \simeq e^{-pN} \frac{(pN)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$$



# Összefüggőség és átmérő

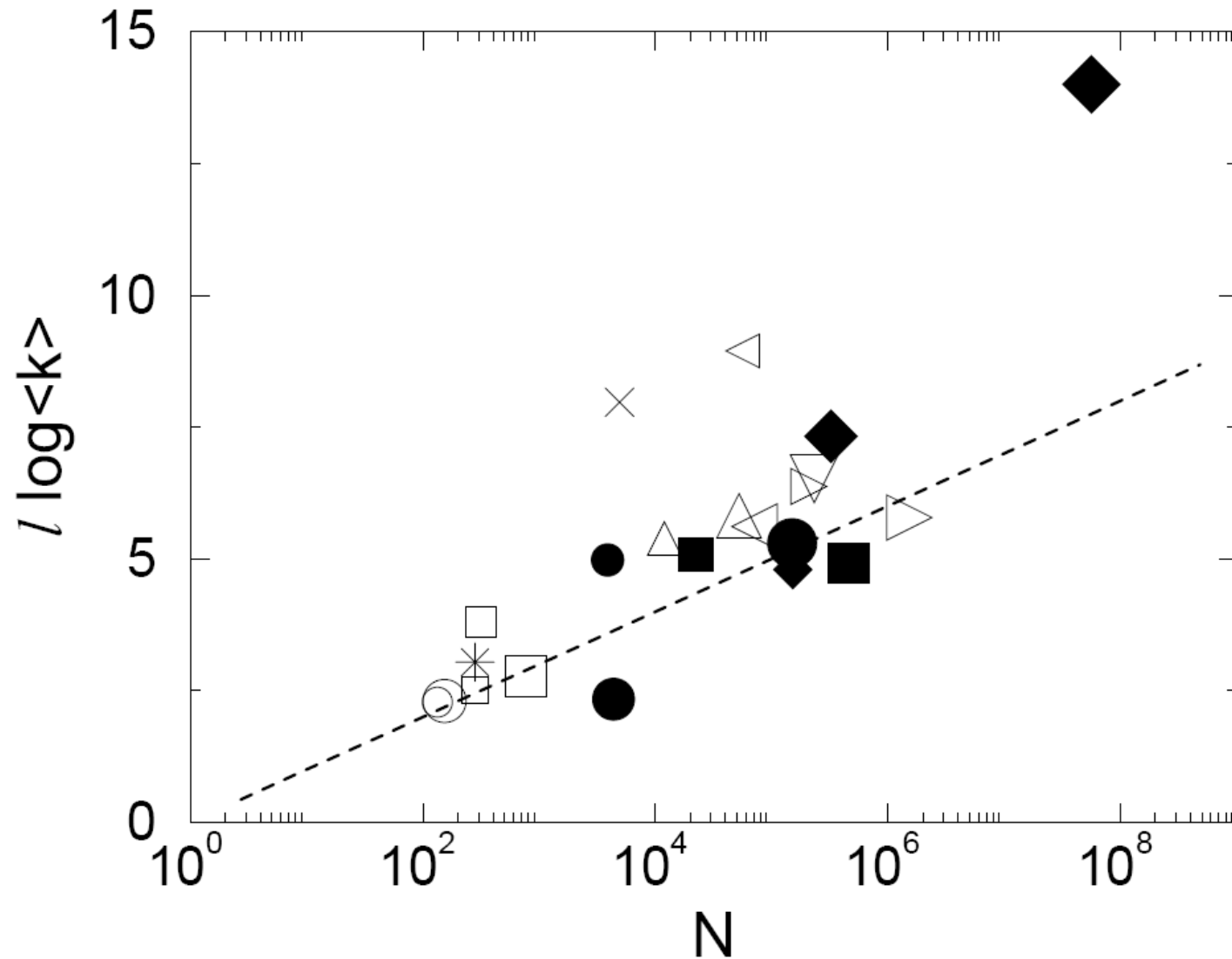
- Ha  $l$  az átmérője a véletlen gráfnak és tudjuk hogy tipikusan  $\langle k \rangle$  a fokszáma akkor  $l$  lépés után érintjük az összes  $\langle k \rangle^l = N$  csomópontot.
- Így tehát  $l = \text{Log}_{\langle k \rangle} N = \frac{\text{Log} N}{\text{Log} \langle k \rangle}$

$$d = \frac{\ln(N)}{\ln(\langle k \rangle)} = \frac{\ln(N)}{\ln(\langle k \rangle)}$$

- Ha  $\langle k \rangle = pN < 1$  esetén kis komponensek  $d$  a legnagyobb komponens átmérője, bár azok tipikusan egyméretűek
- Ha  $\langle k \rangle > 1$  akkor van óriás komponens és annak az átmérőjét mérjük
  - $\langle k \rangle > 3.5$  felett jellemzően arányosnak mondhatjuk  $\text{Log}(N)/\text{Log}(\langle k \rangle)$ -val.
- Ha  $\langle k \rangle \geq \text{Log}(N)$  az  $\ell_{rand} \sim \frac{\ln(N)}{\ln(\langle k \rangle)}$  pontosan  $\text{Log}(N)/\text{Log}(\langle k \rangle)$

Átlagos távolság:

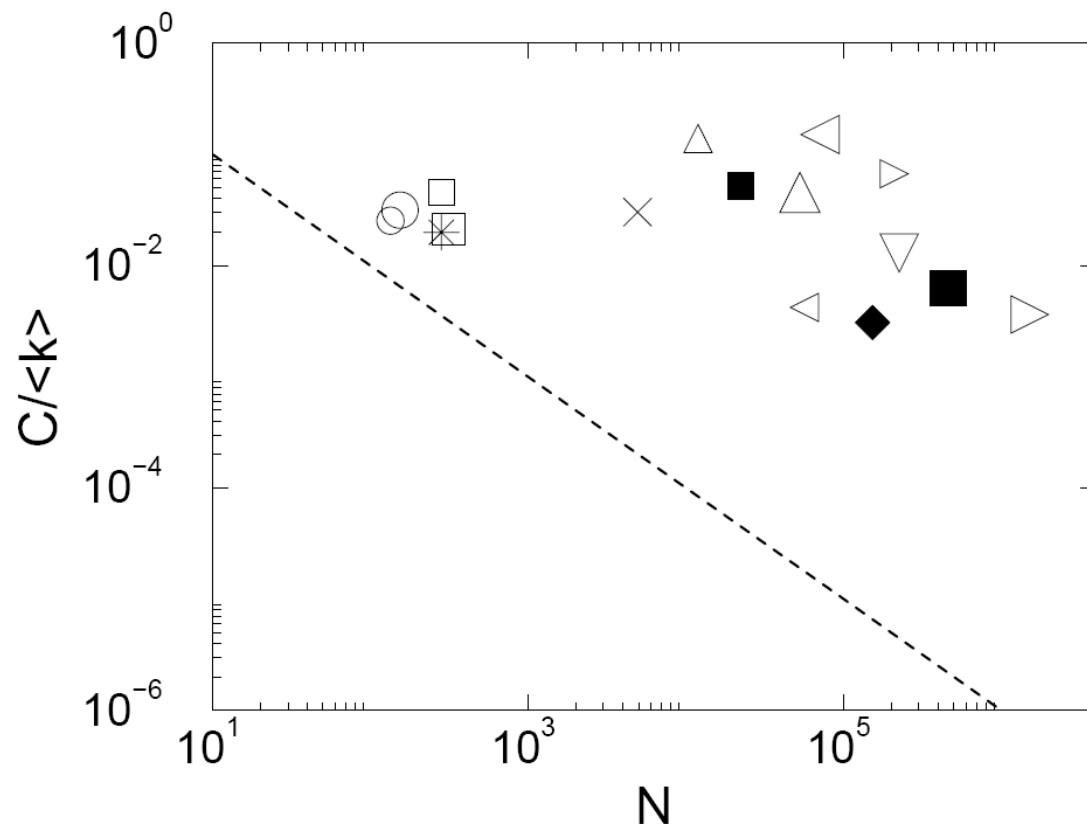
# Átmérő valós és klasszikus véletlen hálózatokban



# Klaszterezési együttható véletlen gráfokban és a valóság

- Mivel minden kapcsolódás függetlenül a többitől  $p$  valószínűséggel következik be így a barátaim barátok-e  $p$  valószínűséggel igaz:

$$C_{rand} = p = \frac{\langle k \rangle}{N}$$



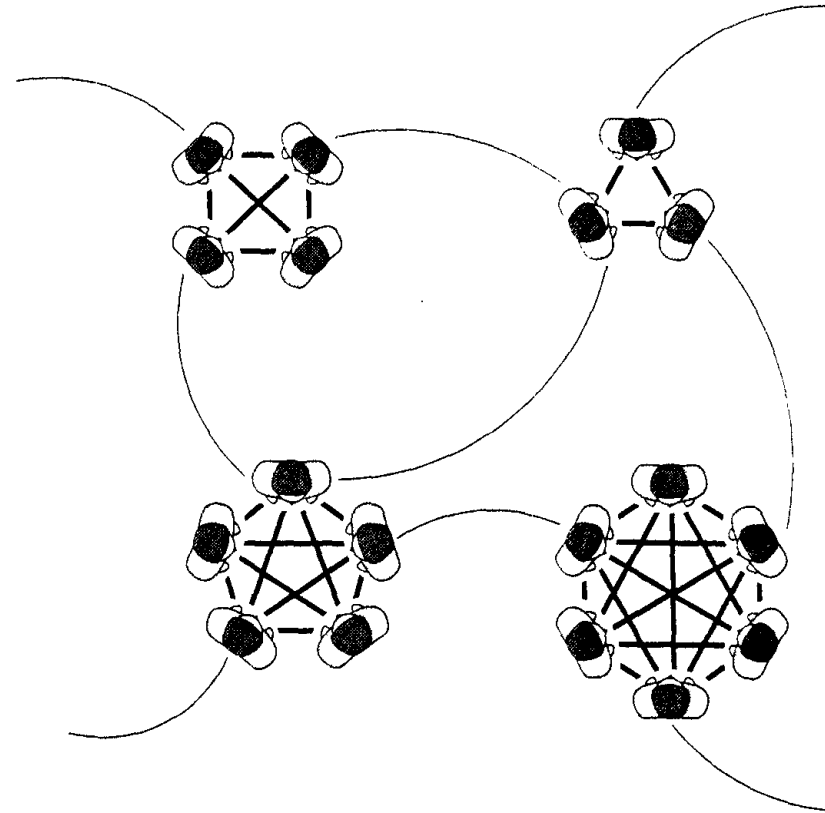
# Áttekintés

	Network or subgraph	Number of vertices	Number of edges	$\gamma$	$C$	$C/C_r$	$\bar{\ell}$	$\bar{\ell}/\bar{\ell}_r$	References	
1	Complete map of the nd.edu domain of the Web	325 729	1 469 680	$\gamma_i = 2.1$ $\gamma_o = 2.4^{\ddagger}$	1	—	—	11.2	—	Albert <i>et al.</i> 1999
2	Pages of the WWW scanned by Altavista <sup>1</sup> in October 1999	$2.711 \times 10^8$	$2.130 \times 10^9$	$\gamma_i = 2.1$ $\gamma_o = 2.7$	2	—	—	16 (6.8)	(1.0)	Kumar <i>et al.</i> 2000a, Broder <i>et al.</i> 2000
3	'————' (another fitting of the same data)	—	—	$\gamma_i = 2.1$ $\gamma_o = 2.8^{\ddagger}$	3	—	—	—	—	Newman <i>et al.</i> 2001
4	Map of sites of a domain of the WWW, spring 1997	$2.60 \times 10^5$	—	$\gamma_i = 1.9^{\ddagger}$	4	—	—	—	—	Adamic <i>et al.</i> 2000
5	Undirected map of sites in a domain of the WWW	153 127	$2.70 \times 10^6$	—	5	0.108	$0.47 \times 10^3$	3.1	0.93	Adamic 1999
6	A set of public company home pages	4923	$1.335 \times 10^7$	$\gamma_i = 2.0^{\ddagger}$	6	—	—	—	—	Pennock <i>et al.</i> 2002
7	A set of US newspaper home pages	—	—	$\gamma_i = 2.0^{\ddagger}$	7	—	—	—	—	Pennock <i>et al.</i> 2002
8	A set of university home pages	—	—	$\gamma_i = 2.6^{\ddagger}$	8	—	—	—	—	Pennock <i>et al.</i> 2002
9	A set of computer scientist home pages	56 880	—	$\gamma_i = 2.6^{\ddagger}$	9	—	—	—	—	Pennock <i>et al.</i> 2002
10	Interdomain level of the Internet, December 1998	4389	8256	2.2	10	—	—	4	0.6	Faloutsos <i>et al.</i> 1999
11	Interdomain level of the Internet, December 1999 <sup>2</sup>	6374	13 641	2.2	11	0.24	$3.3 \times 10^2$	3.7	0.58	Pastor-Satorras <i>et al.</i> 2001
12	Router level of the Internet in 1995	3888	5012	2.5	12	—	—	12.1	1.39	Faloutsos <i>et al.</i> 1999
13	Router level of the Internet in 2000	$\sim 150\,000$	$\sim 200\,000$	$\sim 2.3$	13	—	—	10	0.8	Govindan <i>et al.</i> 2000
14	Citations in the ISI database 1981–June 1997	783 339	6 716 198	$\gamma_i = 3.0$	14	—	—	—	—	Redner 1998
15	'————' (another fitting of the same data)	—	—	$\gamma_i = 2.9$	15	—	—	—	—	Tsallis <i>et al.</i> 2000
16	'————' (another estimate from the same data)	—	—	$\gamma_i = 2.5$	16	—	—	—	—	Krapivsky <i>et al.</i> 2000
17	Citations in <i>Phys. Rev. D</i> 11–50 (1975–1994)	24 296	351 872	$\gamma_i = 3.0$	17	—	—	—	—	Redner 1998
18	'————' (another fitting of the same data)	—	—	$\gamma_i = 2.6$	18	—	—	—	—	Tsallis <i>et al.</i> 2000
19	'————' (another estimate from the same data)	—	—	$\gamma_i = 2.3$	19	—	—	—	—	Krapivsky <i>et al.</i> 2000
20	'————' (another estimate from the same data)	—	—	$\gamma_i = 1.9$	20	—	—	—	—	Vázquez 2001b
21	Collaboration network of movie actors	212 250	61 085 555	2.3	21	—	—	4.54	1.25	Barabási <i>et al.</i> 1999
22	'————' (another fitting of the same data)	—	—	3.1	22	—	—	—	—	Albert <i>et al.</i> 2000a
23	Collaboration network of MEDLINE	1 388 989	$1.028 \times 10^7$	2.5	23	0.066	$6 \times 10^3$	4.6	0.9	Newman 2001e
24	Coauthorships in the SPIRES <sup>3</sup> e-archive	56 627	$4.898 \times 10^6$	1.2	24	0.726	$0.24 \times 10^3$	4.0	1.88	Newman 2001e
25	Collaboration net collected from math. journals	70 975	$0.132 \times 10^6$	2.1	25	0.59	$1.1 \times 10^4$	9.5	1.16	Barabási <i>et al.</i> 2002
26	Collaboration net collected from neurosci. journals	209 293	$1.214 \times 10^6$	2.4	26	0.76	$1.4 \times 10^4$	6	1.2	Barabási <i>et al.</i> 2002
27	Web of human sexual contacts <sup>4</sup>	2810	—	3.4	27	—	—	—	—	Liljeros <i>et al.</i> 2001
28	Networks of metabolic reactions <sup>5</sup>	$\sim 200\text{--}800$	$\sim 600\text{--}3000$	$\gamma_i = 2.2$ $\gamma_o = 2.2$	28	0.32	12	3.2	0.95	Jeong <i>et al.</i> 2000, Wagner <i>et al.</i> 2001
29	Net of protein–protein interactions (yeast proteome) <sup>6</sup>	1870	2240	$\sim 2.5$	29	0.022	4.4	6.8	0.8	Jeong <i>et al.</i> 2001, Wagner 2001a
30	Word Web <sup>7</sup>	470 000	17 000 000	1.5 / 2.7	30	0.69/0.44	$4.4 \times 10^3 / 2.8 \times 10^3$	2.65	0.87	Ferrer i Cancho <i>et al.</i> 2001a
31	Food web of Silwood park <sup>8</sup>	154	366	$\sim 1$	31	0.15	5	3.4	1.05	Montoya <i>et al.</i> 2001, Solé <i>et al.</i> 2002
32	Java Development Framework (largest component)	1376	2174	2.5	32	0.06	25	6.39	1.02	Valverde <i>et al.</i> 2002
	'————' (second largest component)	1364	1930	2.65	—	0.08	40	6.91	1.01	Valverde <i>et al.</i> 2002
33	Computer game graph	1989	$4.78 \times 10^3$	2.85	33	0.08	35	6.2	1.28	Valverde <i>et al.</i> 2002
34	Large digital electronic circuits	$2 \times 10^4$	$4 \times 10^4$	3.0	34	$3 \times 10^{-2}$	$1.5 \times 10^2$	$\sim 6$	$\sim 1$	Ferrer i Cancho <i>et al.</i> 2001b
35	Telephone call graph <sup>9</sup>	$47 \times 10^6$	$8 \times 10^7$	$\gamma_i = 2.1$	35	—	—	—	—	Aiello <i>et al.</i> 2000
36	E-mail net (student accounts in Kiel University) <sup>10</sup>	5165	$6.57 \times 10^4$	1.3, $\gamma_i=1.5$	36	0.156	$3.25 \times 10^3$	4.95	0.48	Ebel <i>et al.</i> 2002
37	Energy landscape network for a 14-atom cluster	4196	87 219	2.78	37	0.073	7.4	2.32	1.04	Doye 2002



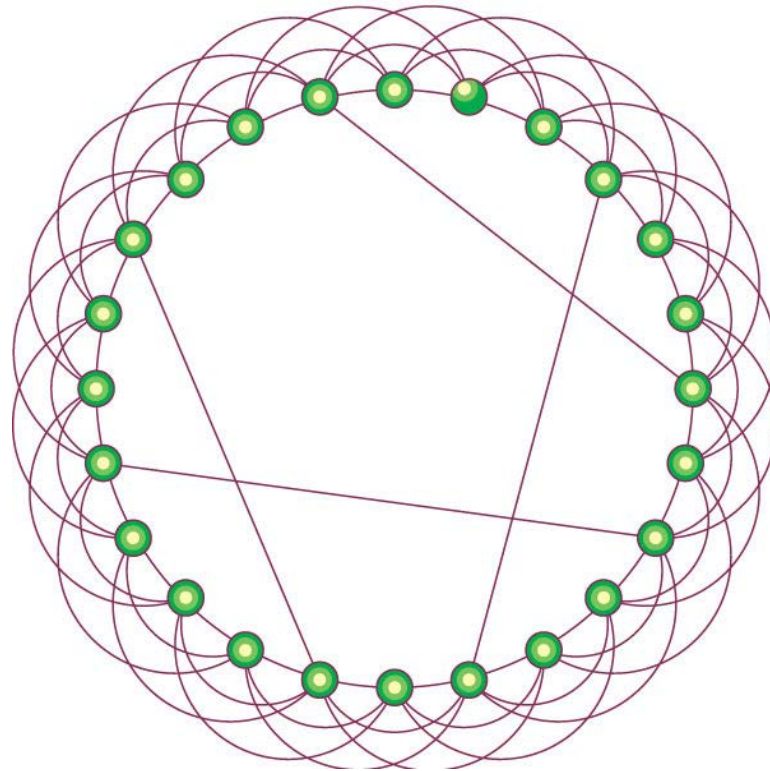
# A gyenge kapcsolatok ereje

- Mark Granovetter: Álláskeresés → A társadalom szerkezete
- Magas klaszterezettség → Nem random
- Távoli kapcsolatok! Különben az átmérő nem lesz kicsi.



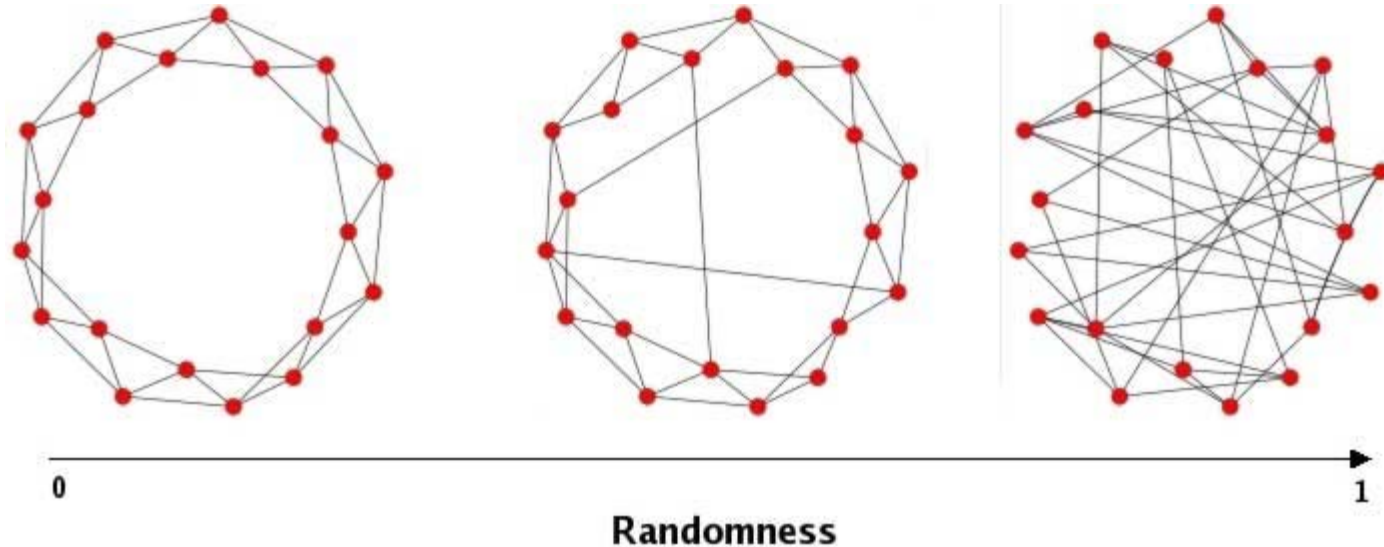
# Gyenge kapcsolatok általános modell

- Két rész:
  - Körvonalon elosztva csomópontok, mindenkinek  $k$  távolságig minden szomszéddal kapcsolat: rövidtávú kapcsolat – klaszterezettség
  - Véletlenszerű távoli kapcsolatok: kis átmérő

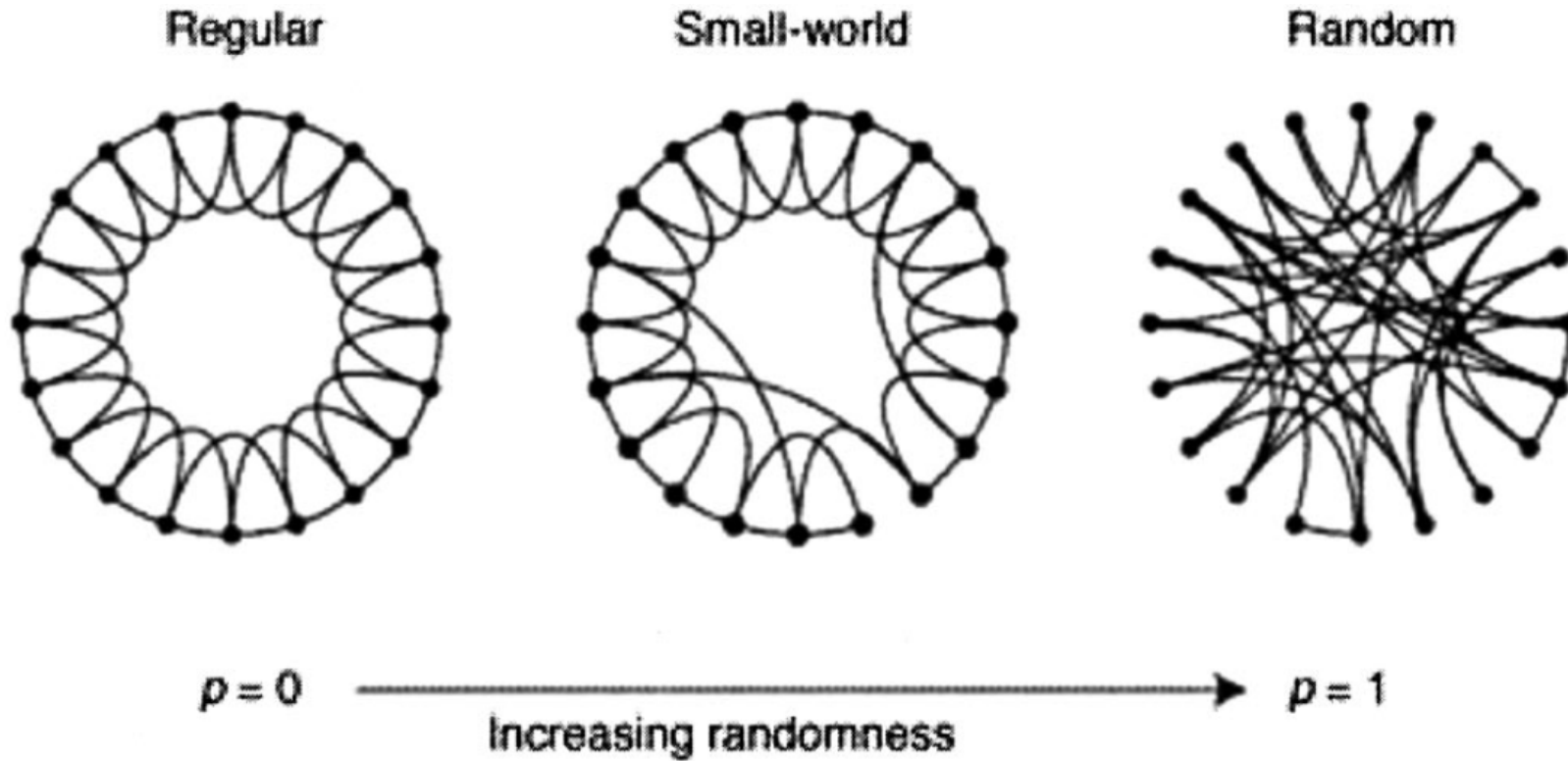


# Véletlen kapcsolatok aránya

- A véletlen kapcsoltok növelésével lassan visszakapjuk az Erdős féle véletlen gráfokat



# Kisvilág modell – Watts an Strogatz



# Kisvilág modell paraméterei

- Klaszterezettség: két távolságban levő szomszédok esetén (előző fólia esete):

$$C=4/6$$

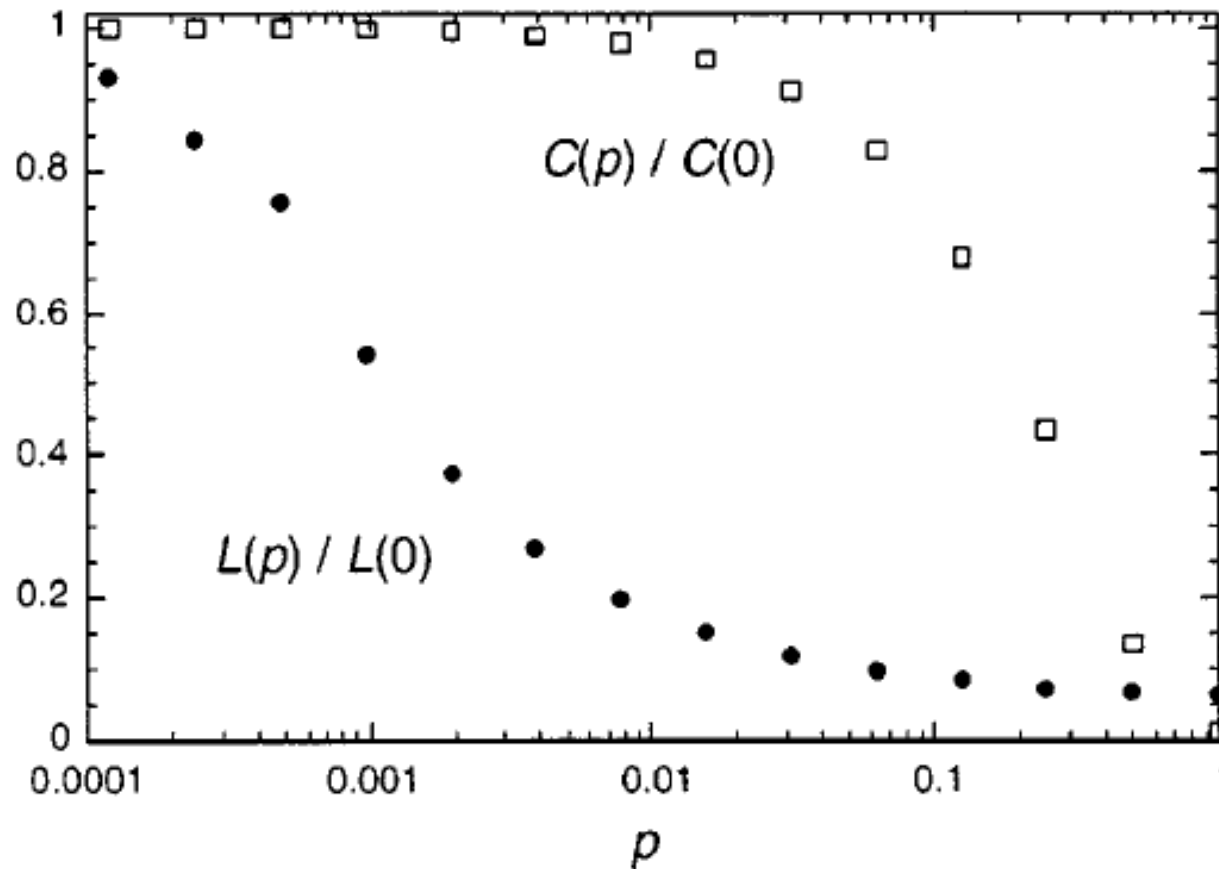
- Ha  $K$  legközelebbi szomszédal van kapcsolatban akkor:

$$C = \frac{3(K-2)}{4(K-1)}$$

- $K \rightarrow$  végtelenben a  $C \rightarrow 0.75$

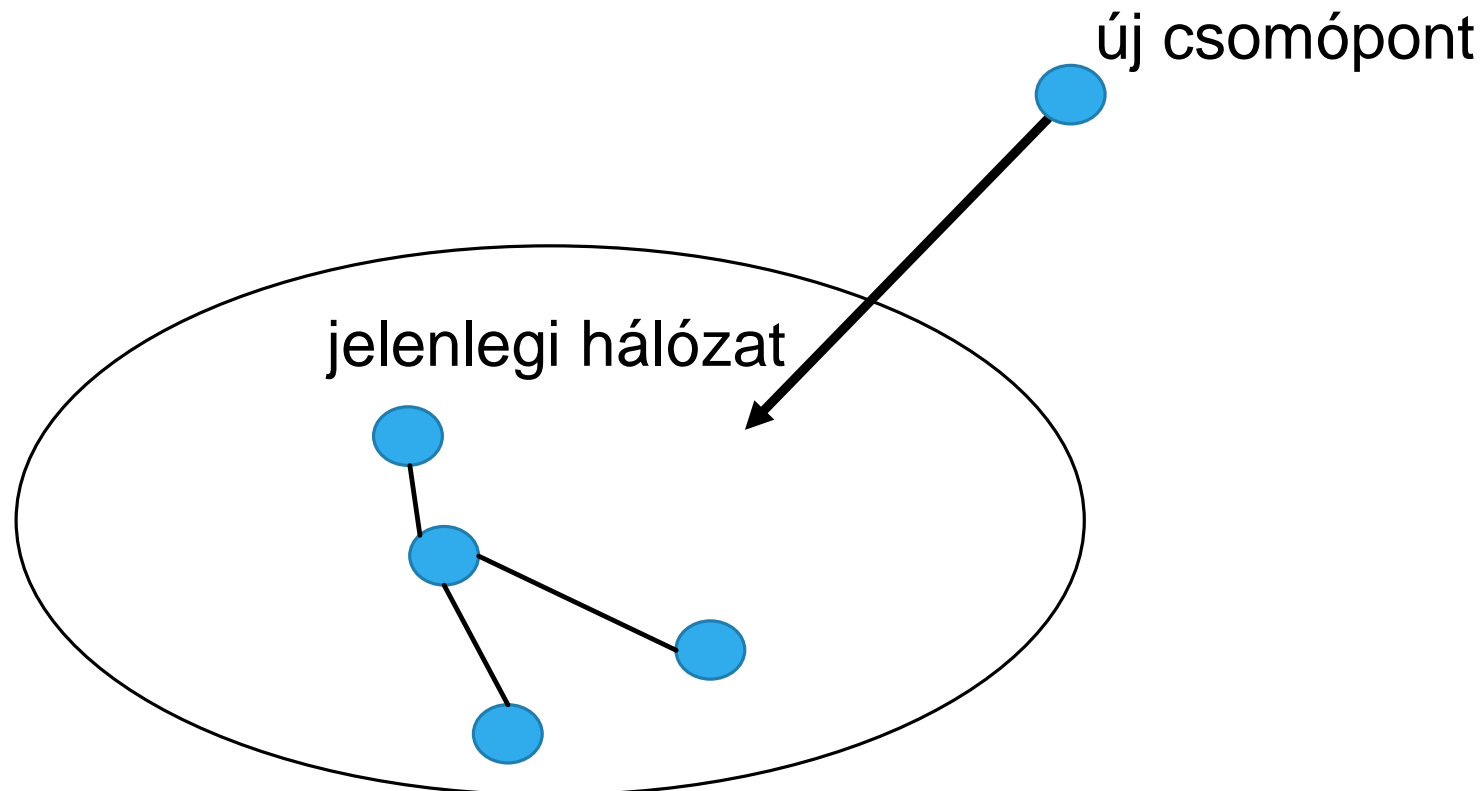
# Watts-Strogatz modell – általános észrevétel

- Minden csomópont  $K/2$  szomszédal mindkét oldalról, mondjuk:  
 $N \gg K \gg \ln(N) \gg 1$
- Húzzunk újra  $p$  valószínűséggel néhány élet:  $pN \ll K/2$  él



# B-A modell

- Barabási-Albert modell
- Megoldás: hálózatevolúció, azaz a hálózat lépésenként alakul



Az új csomópont a legnagyobbhoz csatlakozik a legvalószínűbben:  
A gazdag csomópont még gazdagabb lesz!





# **Folyamatok nagy hálózatokon**

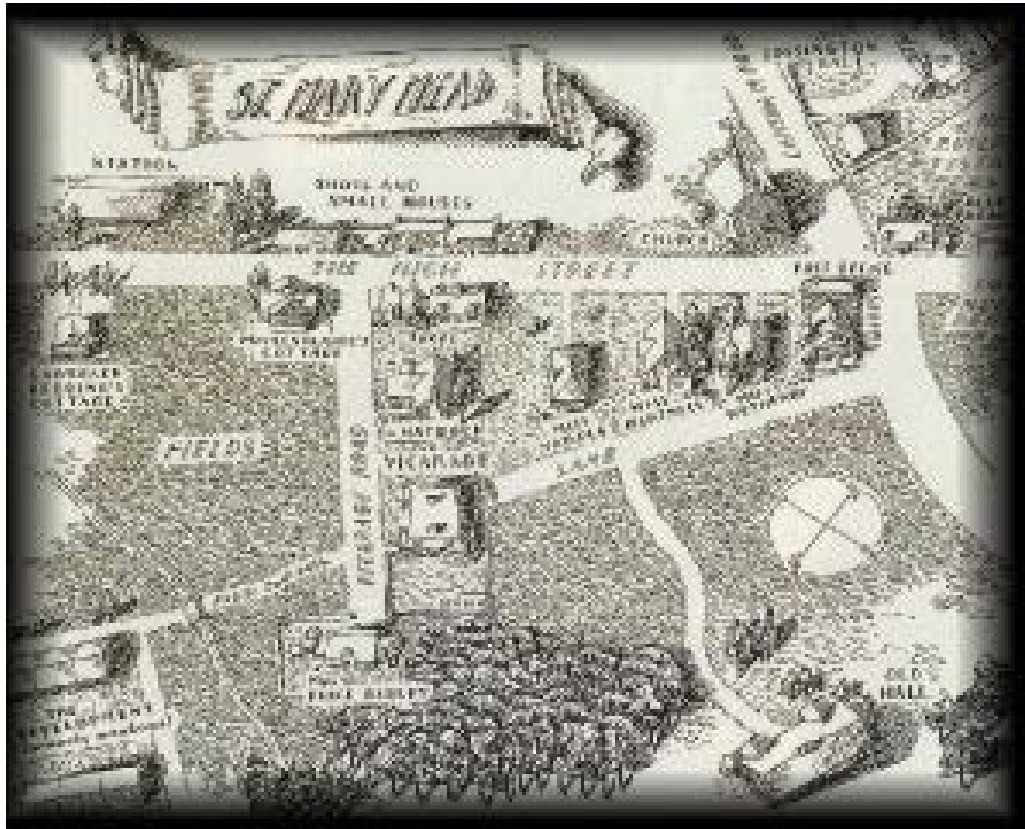
## **Keresés az Interneten**

# Mi a közös?



# Mi köze a Miss Marple-nek a hálózattudományhoz?

- Marple kis faluja: St. Mary Mead

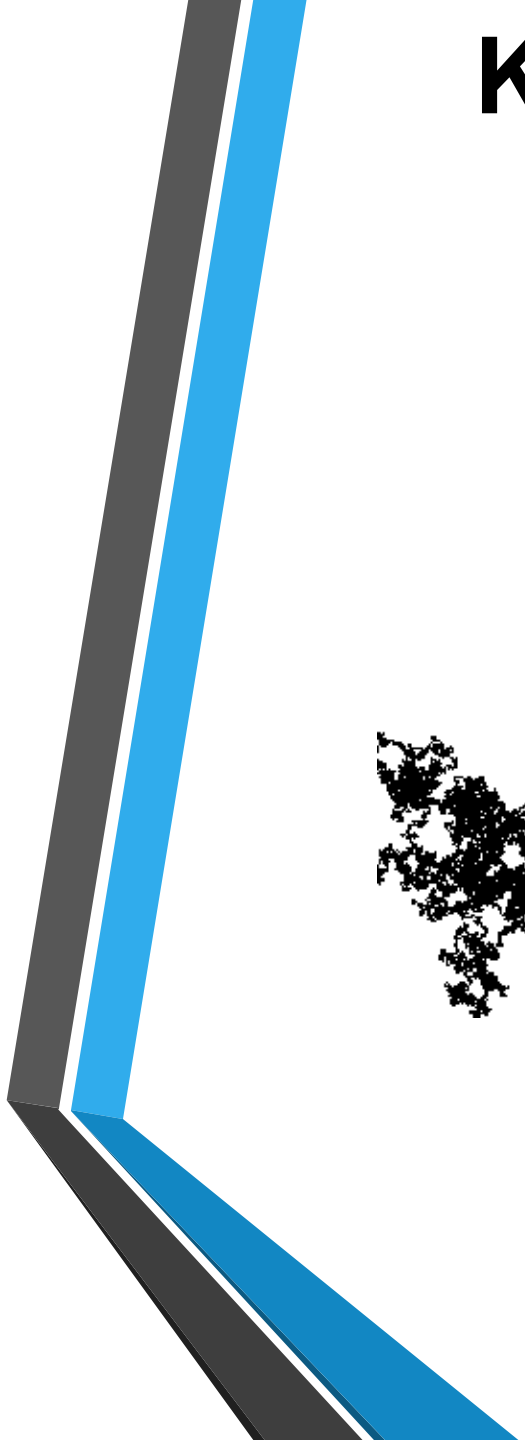
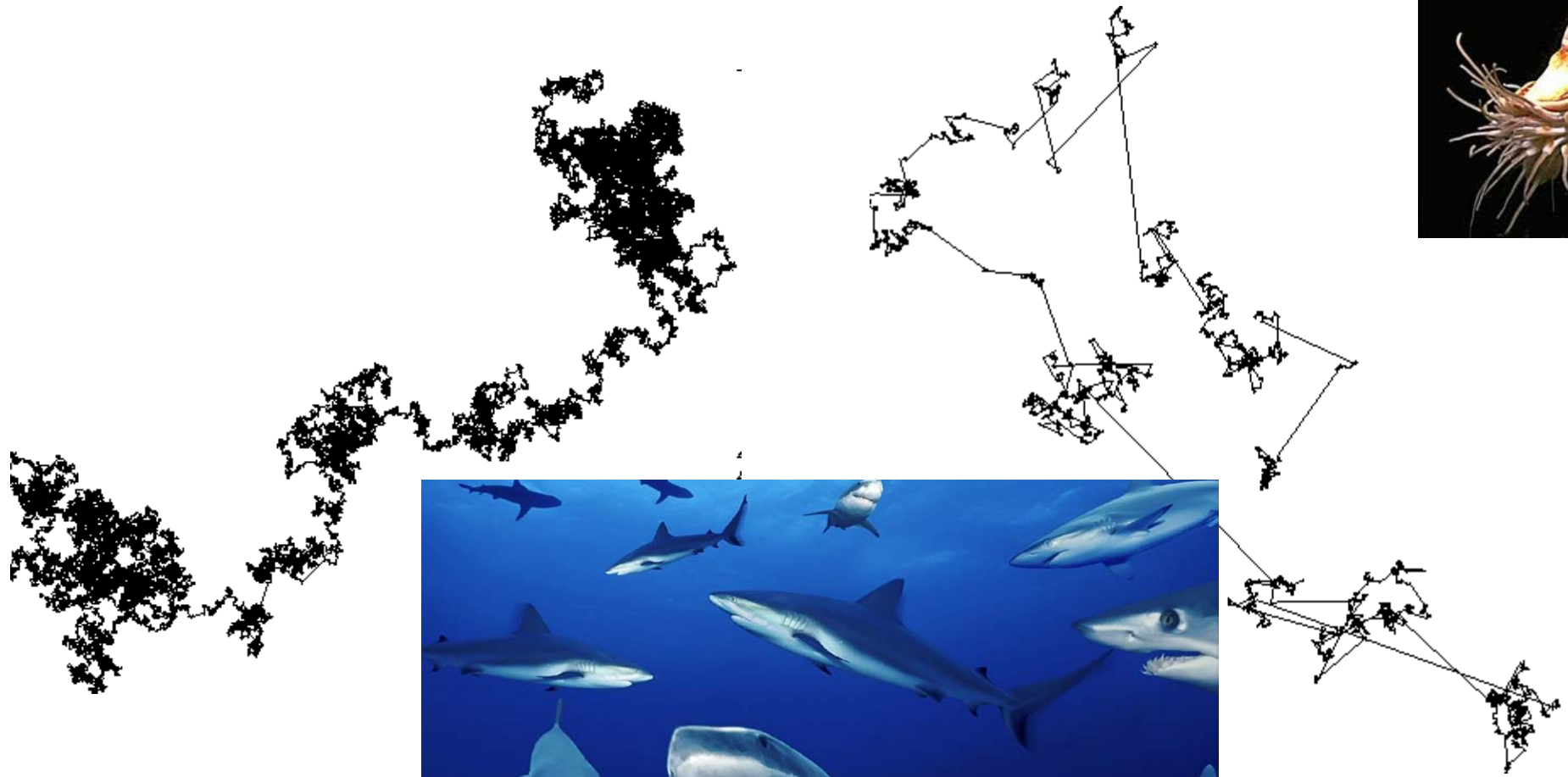


# Keresés

- Az élet egy nagy keresés 😊
  - kaját
  - megoldást
  - választ
  - erőforrást
  - Lehetőséget
- Általában térkép nélkül
- Teljes/elárasztott keresés vagy
- Irányított keresés

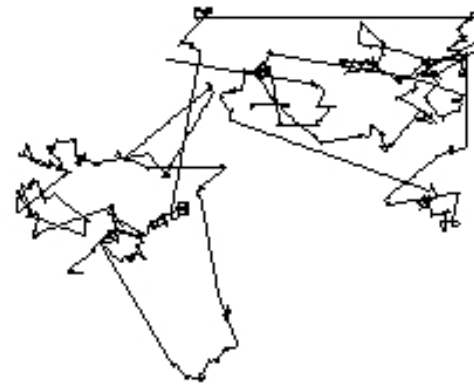


# Keresési stratégiák és az evolúció

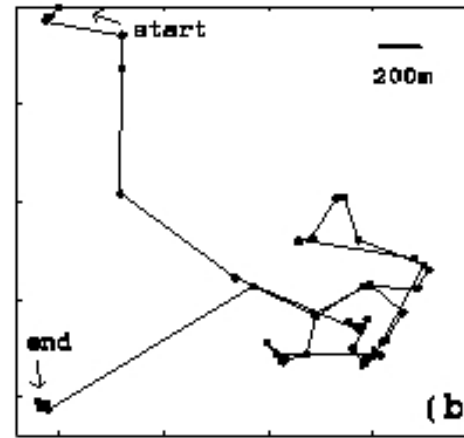




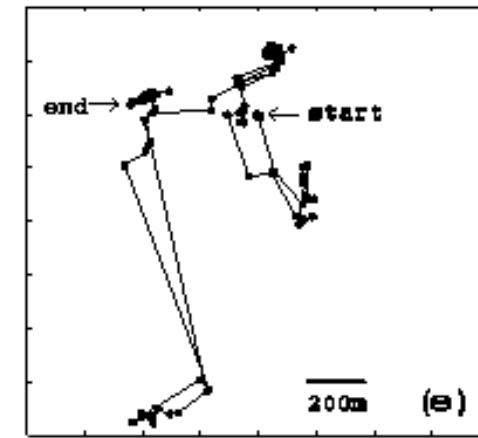
# Emberek téblábolnak a vidámparkokban...



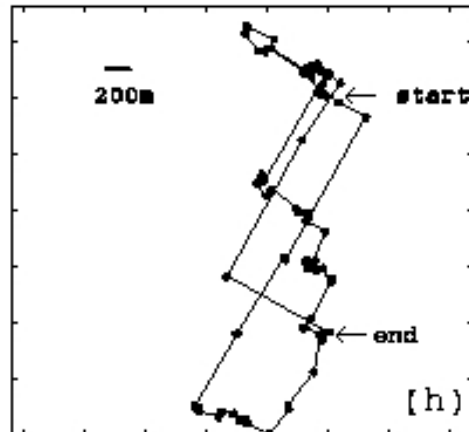
Levy Walks  
(randomly generate)



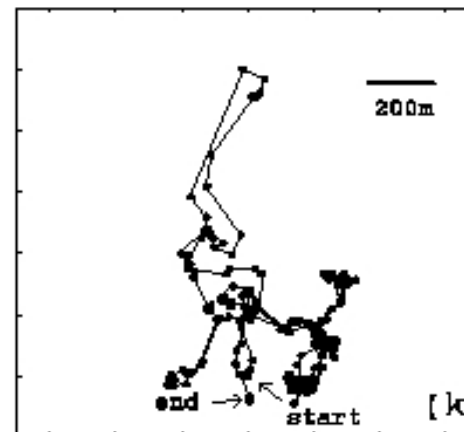
NCSU



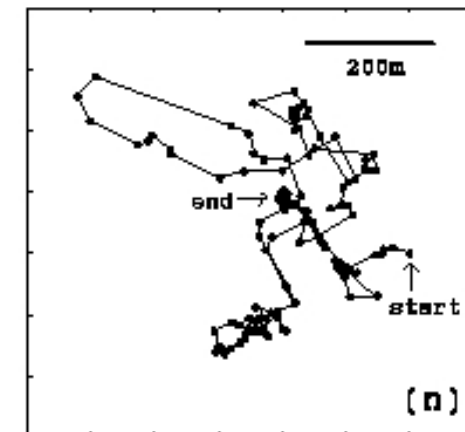
KAIST



NYC (Manhattan)



Disney World



State Fair

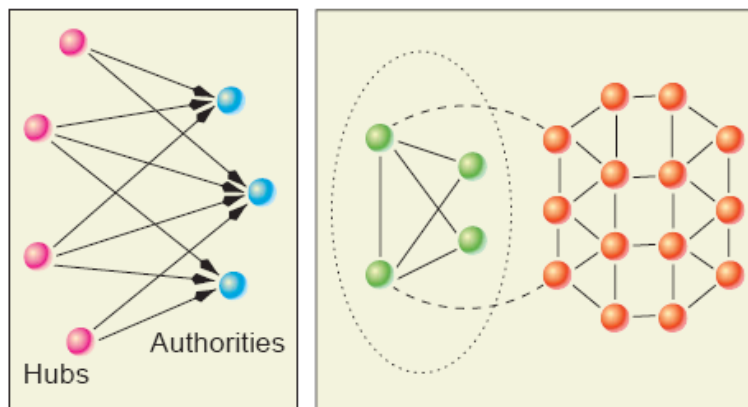
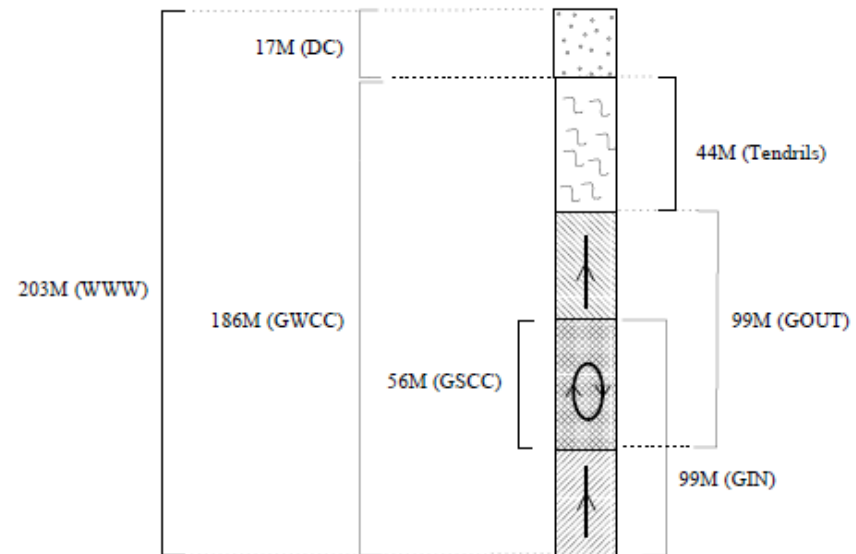
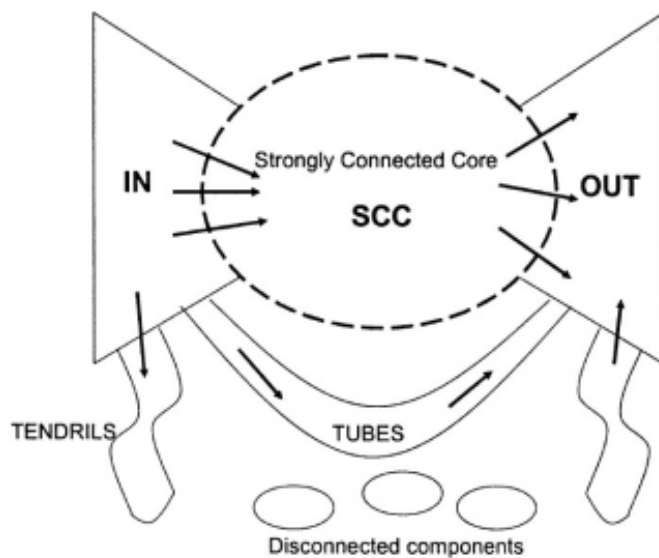
# Keresés a weben

- Szörfölés
- Levy utak mentén
- Robotokkal (szélességi bejárás)
  - crawler, spider
  - Indexelés
  - Találatok rangsorolása
- A nagy keresők vajon hány százalékát ismerik a webnek?
- Milyen a web szerkezete?
- A "long tail modell" a weben...
  - web 2.0?
- Web 3.0??
  - Geoweb?





# A web szerkezete



Kereshető web kevesebb mint 50%

Irányított út valószínűsége <25%

Keresők kapacitása nő

De a Web gyorsabban

Keresők evolúciója

- Vizsgálat célja a pusztán információ elérésén kívül:  
Társadalom, tartalom tanulmányozása, reklámozás, stb.

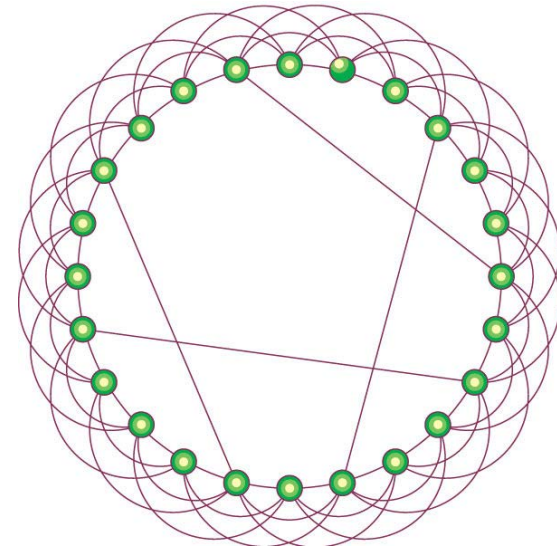
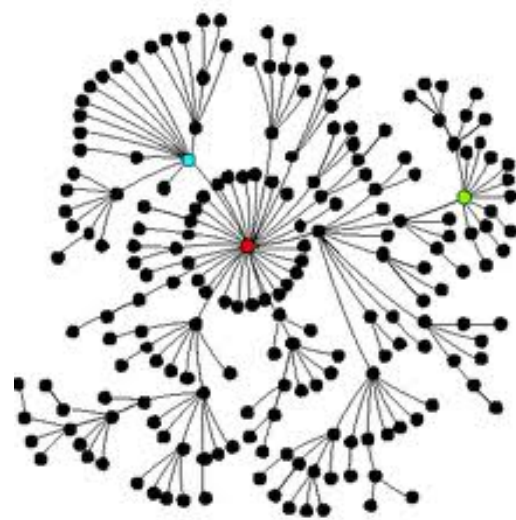
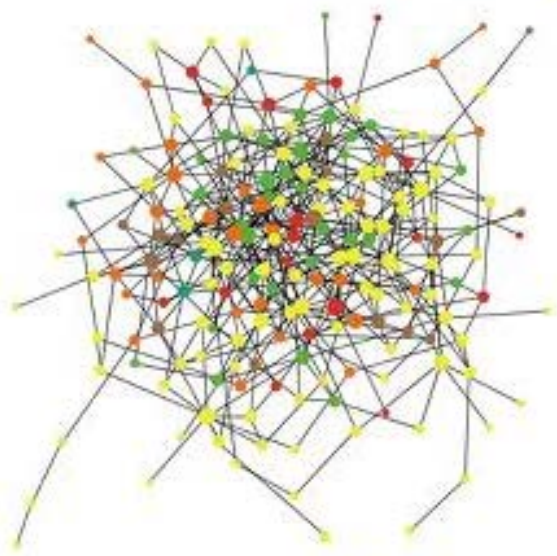
# Keresés, kereshető hálózatok

- Már megint Milgram
- Lehetne kicsi a világ a kísérlet sikertelensége ellenére is
- Rövid utak, és meg is találjuk őket térkép nélkül
- Elárasztásos keresés
- Erdős szám meghatározása gép nélkül
- Véletlen bejárás, DS bejárás mint a Gnutellánál
- De ahhoz erős hubok kellene
- Irányított keresés:
  - Legjobb szándék szerint elindítjuk (De mi alapján?)
  - De biztosan jó irányba megy?
- Irányított keresés példák:
  - Szörfölés, p2p filekeresés, üzleti kapcsolat keresés, erőforrás keresés, megoldás keresése a problémánkra



# Kereshető hálózatok

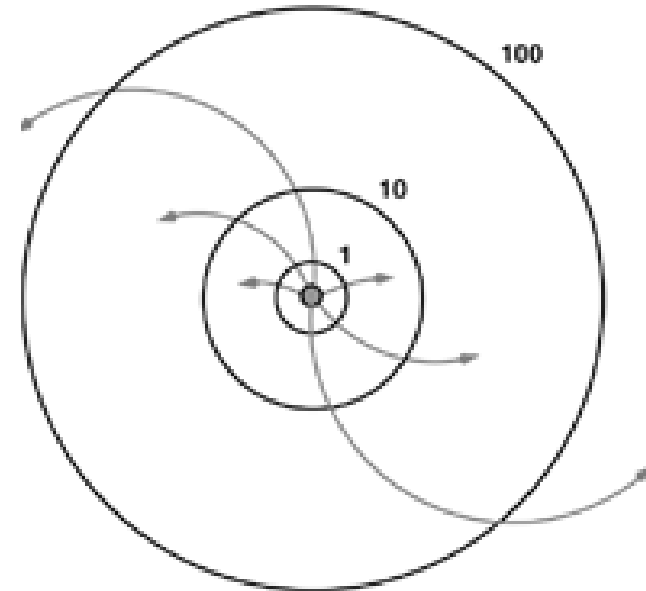
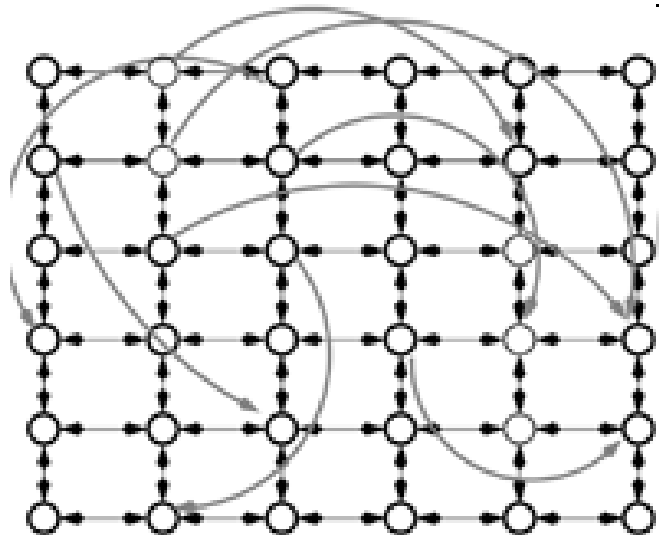
- Véletlen gráf kereshető?
- B-A modell kereshető?
- W-S modell kereshető?



- Mi hiányzik?

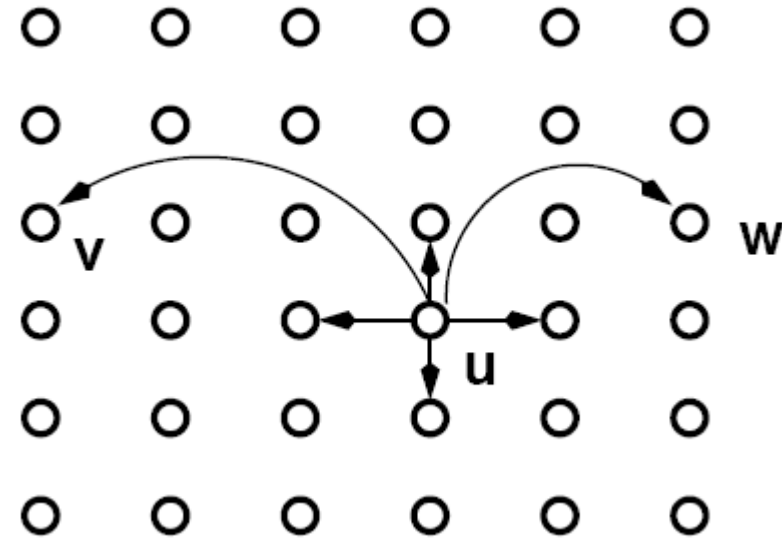
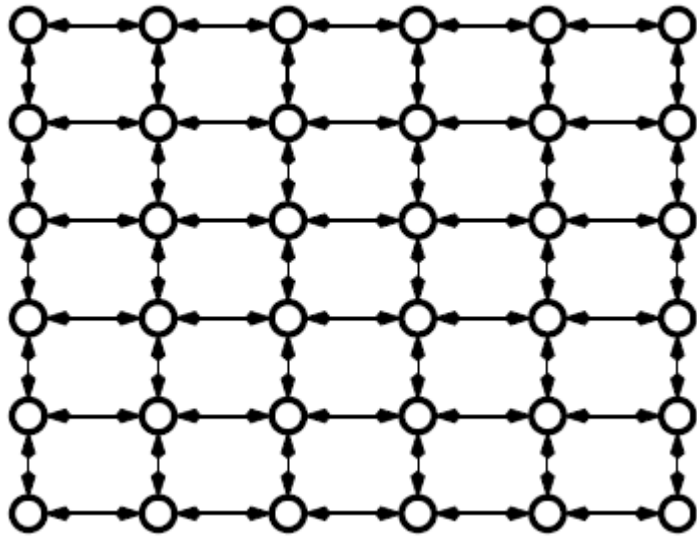
# A távolságfogalom

- Kleinberg modellje
  - Hosszútávú kapcsolat valószínűsége a távolságtól függően



# A célkeresési algoritmus

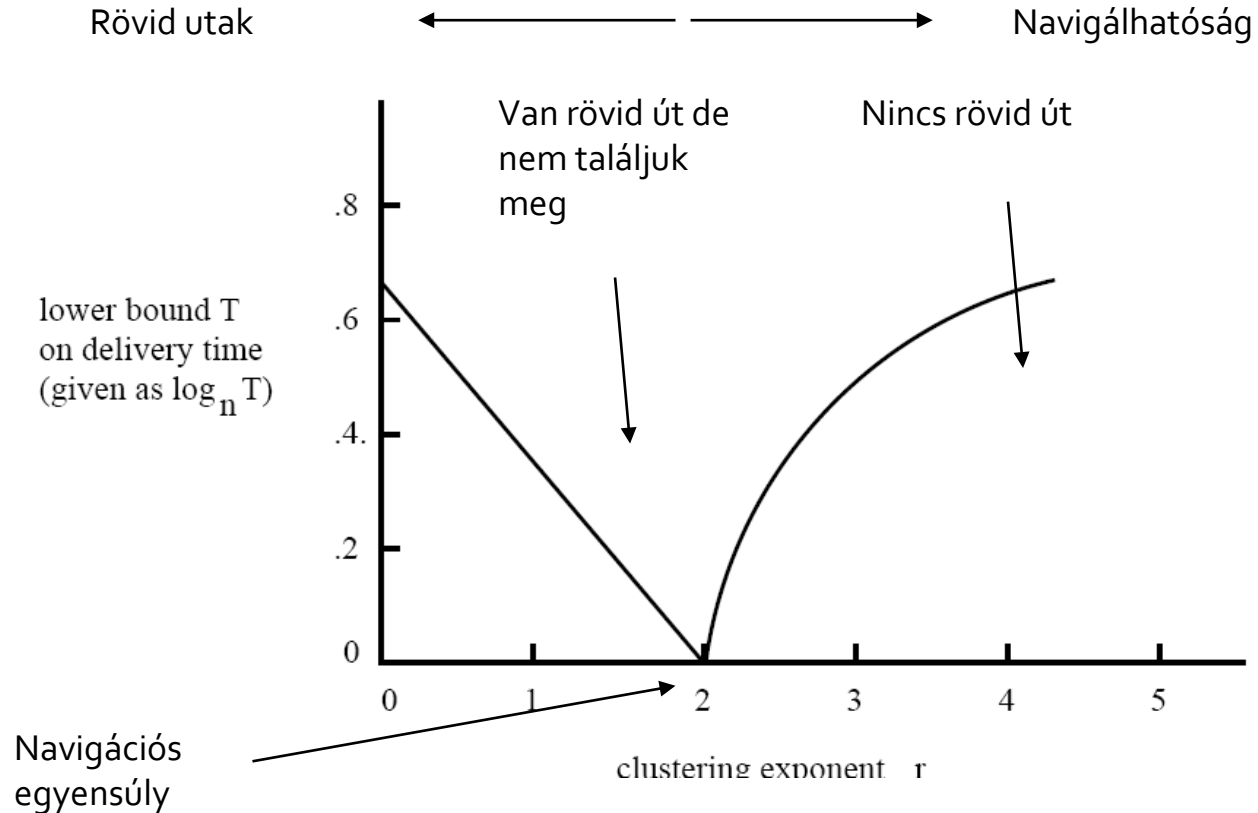
- Jon Kleinberg: Nem csak a topológia érdekes, hanem hogy gyorsan meg is lehet találni a célt, térkép nélkül
- Az optimális modell kereséséhez



- Távolság:  $d(u,v) \rightarrow$  lépések száma a szomszédokon
- A rácson két pont között az kapcsolat valószínűsége  $\sim d(u,v)^{-r}$
- Mohó keresési algoritmus

# Az optimiális topológia

- A rácson két pont között az kapcsolat valószínűsége  $\sim d(u,v)^{-r}$

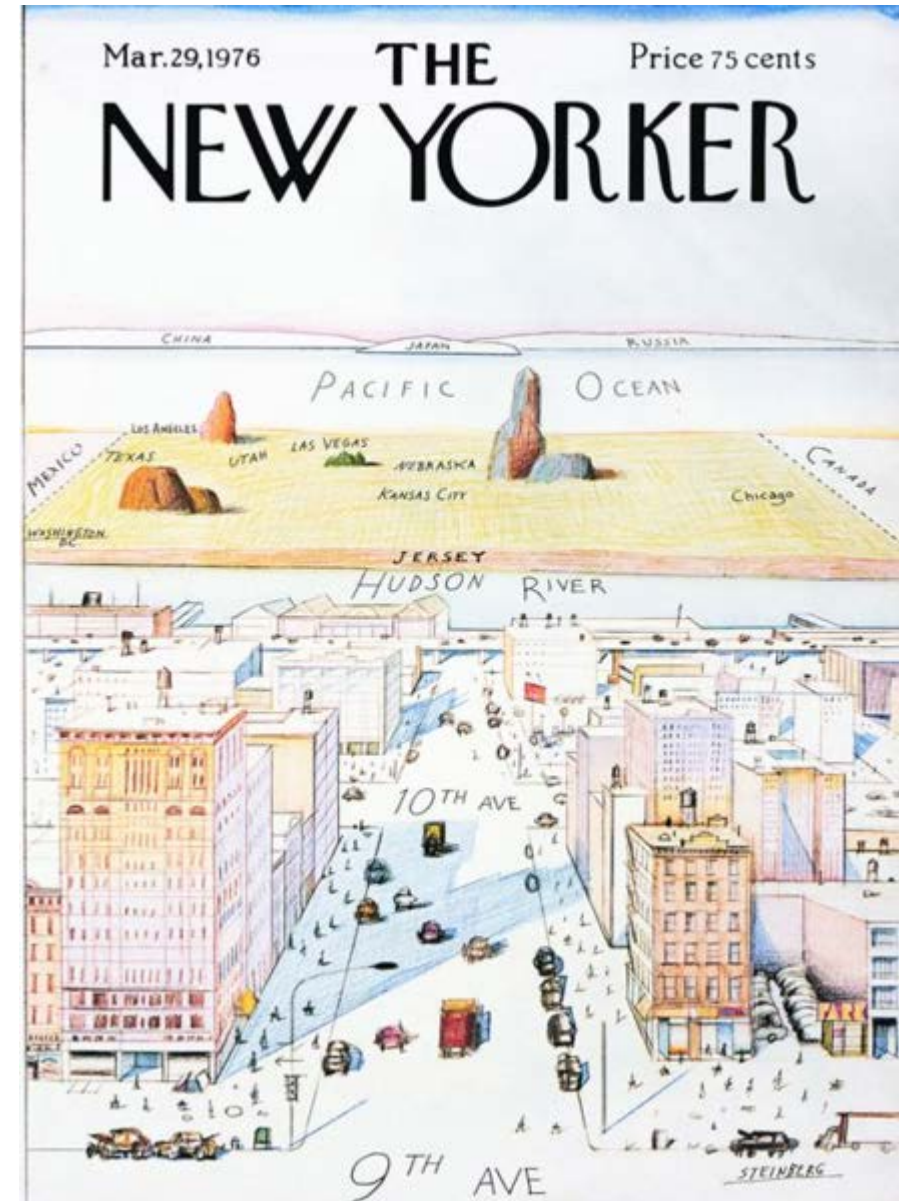
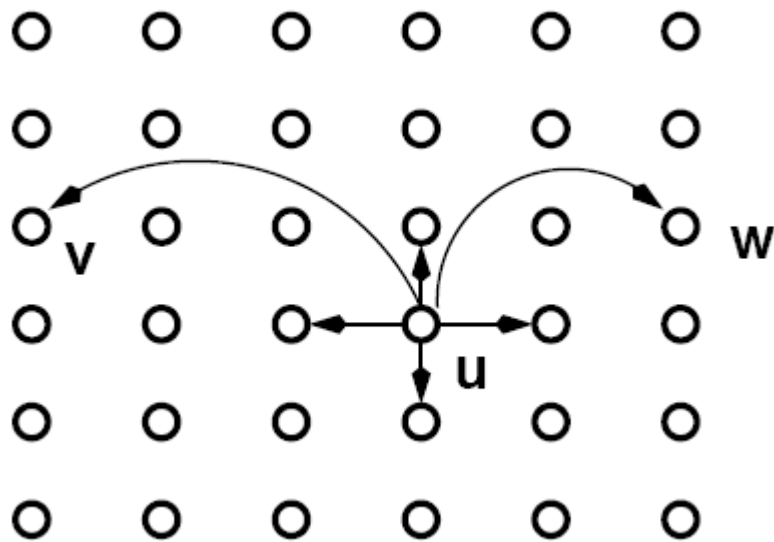


• Így vagy összevissza ugrálunk, vagy lassan haladunk a „rövid” kapcsolatok mentén



# Konceptió

- Minden skálán ugyanannyi kapcsolat van
- A rácson két pont között az kapcsolat valószínűsége  $\sim d(u,v)^{-r}$



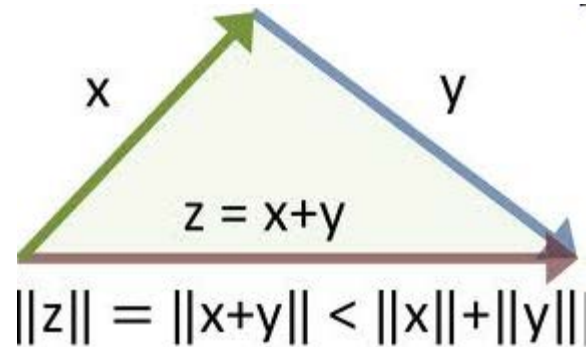


# A Kleinberg modell értékelése (Kleinberg 2002)

- Nemcsak léteznek az utak
- De könnyen megtalálhatók
  - Csak továbbítani kell annak aki a legközelebbinek látszik
  - (Először a megfelelő kontinensre, országba, aztán megyébe stb.)
  - Az ő nézőpontjából a hálózat pontosan ugyanúgy néz ki
  - De ő több információval rendelkezik azon a „környéken”
- A kereséshez nem elég csupán a véletlen shortcut
- Ahhoz, hogy használni lehessen őket, információt kell kódolniuk az adott struktúráról (p2p finger tábla)
- Klaszterezettség kell
- De hogyan lesz ilyen speciális a világ?  
Magyarázza ez a Milgram kísérletet?

# Szociális távolság alapú navigálás (Watts 2003)

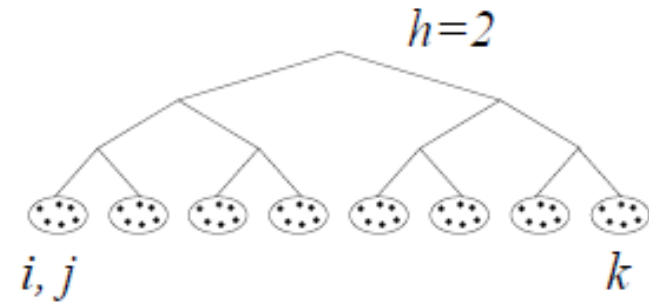
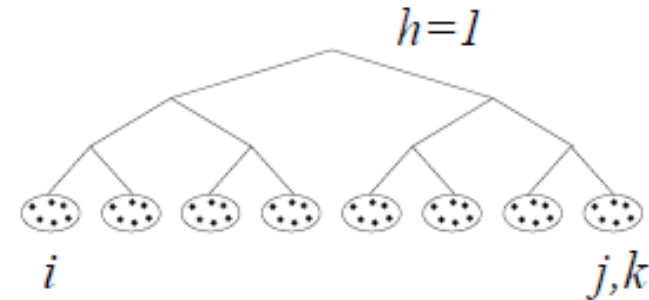
- Hogy lesz két kis ugrásból egy nagy?



1. Csoportosítjuk a többieket
2. A szociális térben elrendezzük őket egy távolság alapján
3. Legrövidebb távolság a dimenziók között

- Távolság a legkisebb közös ősz szintje a hierarchiában
- Minél távolabb vannak, annál kisebb az esély, hogy ismerjük (homofília) mint Kleinbergnél

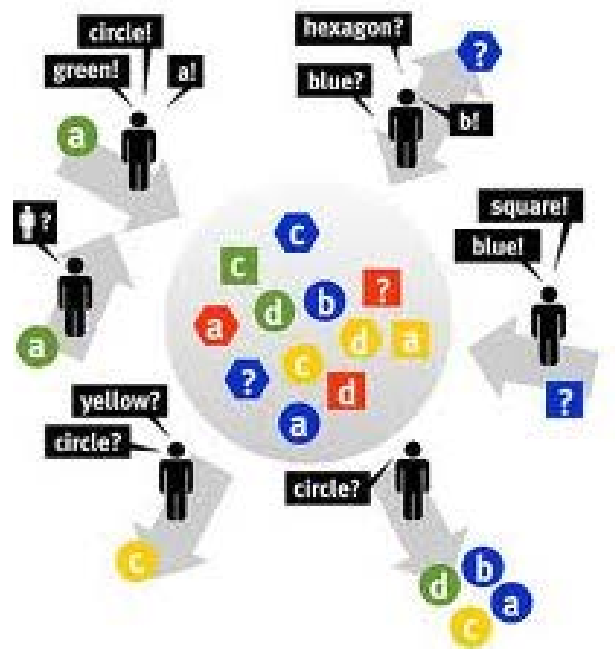
Több szociális dimenzió  
(faj, szakma, vallás, nyelv, kor, külső stb.)



$$y_{ik} = 4 > y_{ij} + y_{jk} = 1 + 1 = 2.$$

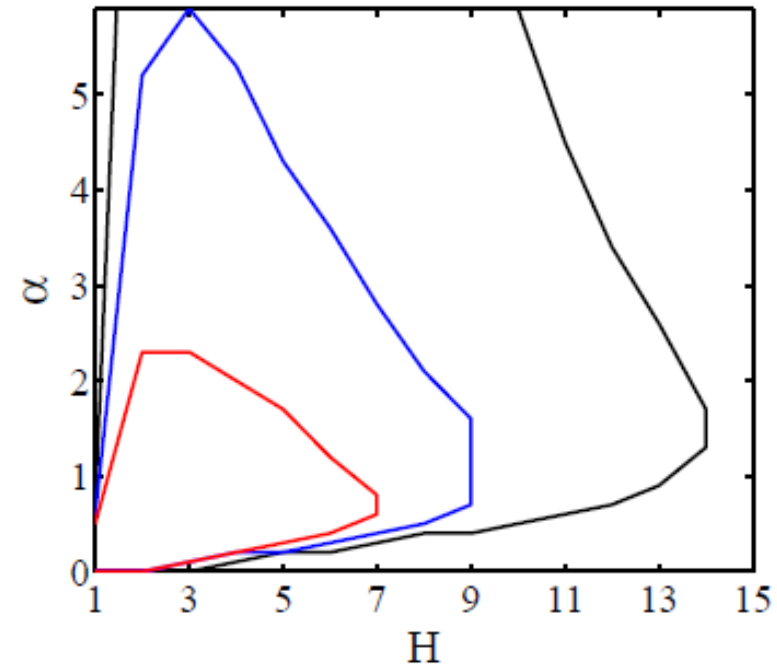
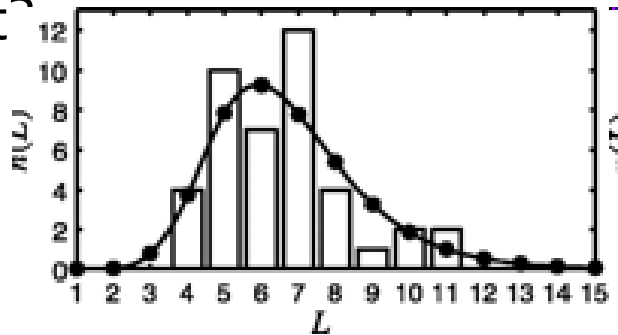
# Szociális dimenziók

- Vajon hány ilyen dimenzió mentén gondolkodunk?
  - Szociológusok szerint max. 5-6
  - Emberre jellemző kognitív határ → Növekvő hálózat miatt átrendeződés



# Szociális távolság alapú navigálás értékelés

- Kereshető hálózatok
- Mohó keresési algoritmus
- Kell hozzá homofília és több mint egy dimenzió
- Eredmények azt mutatják, hogy általában csak 2-3-at célszerű használnunk a lehetséges kb. 6-ból)
  - Földrajzi távolság
  - Szakma
- Miért?



# Keresés (Simsek és Jensen 2008)

- Azt választjuk ahol minimális a várható úthossz:

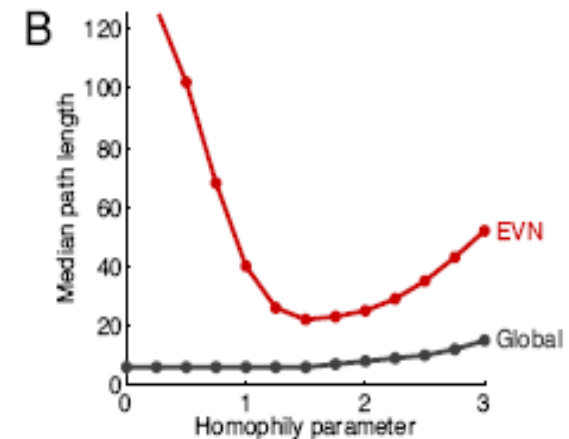
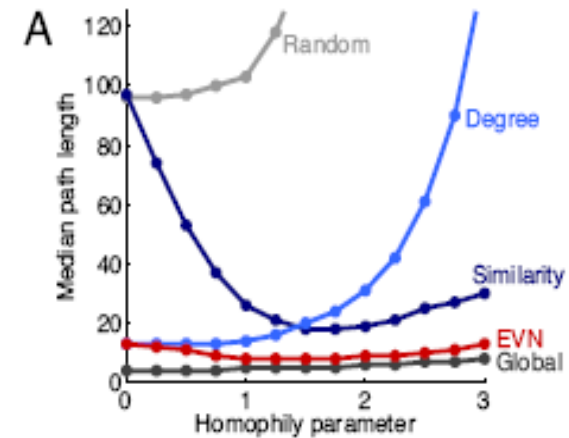
$$E(d_{st}) = \sum_{i \geq 1} i \cdot P(d_{st} = i).$$

- Ehhez maximalizáljuk  $k_s f_{s,t}$ -t ahol

$$f_{s,t} = (x_s - x_t)^{-\alpha}$$

- Hát nem szép ez?
  - Véletlenség  $\rightarrow$  rövid utak
  - Homofília  $\leftrightarrow$  klaszterezettség
  - Skálafüggetlenség

- A: Skálafüggetlen gráf 1.5 paraméterrel
- B: Véletleg gráf 3.5 átlagos kapcsolattal



# P2P keresési technikák

- Problémák:

- Olcsó megoldás
- Támadhatóság
- Sebesség
- Menedzselhetőség

- Folyamatosan fejlődő technikák

- Napster
- BitTorrent
- Reverse sharing

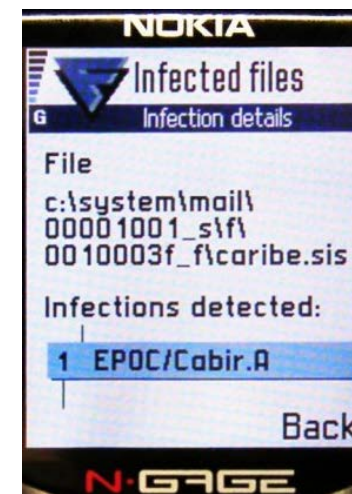
- Vajon alkalmazható ebben a világban amit a társadalmi hálókról megtudtunk?

- Metrikus tér keresése
- Mesterséges beágyazás



# Vírusok és egyébek

- HIV, Ebola, Influenza
  - Fertőzési tulajdonságok+hálózat
  - Afrikai esőerdők, kis mozgástér (lappangási idő alatt)
  - De ma szinte korlátlanok az utazási lehetőségek
  - Ebből a szempontból:
    - Kisvilág ☹️
    - Skálafüggetlenség ☹️ (Gaetan Dugas)
    - Klaszterezettség: 😊
- Számítógép vírusok
  - Internet előtt (floppy-n)
  - Az Internet elterjedésével nulla energiával
  - Broadcast keresés (mindegy kit)
  - Exponenciális növekedés
  - Melissa, Klez, Bugbear, Sobig, Mydoom, Netsky, Bagle
  - Ma már inkább észrevétlenség, adatszerzés, kapacitás
  - Cabir 2004 (mobil bluetooth)





# Vírusok jellegzetességei

- Vírus
  - utasításhalmaz ami elsősorban önmaga sokszorosításáról szól
- Mennyire fertőző
- Mennyi ideig tartja a gazdát fertőző állapotban

- Vírusterjedés vizsgálata



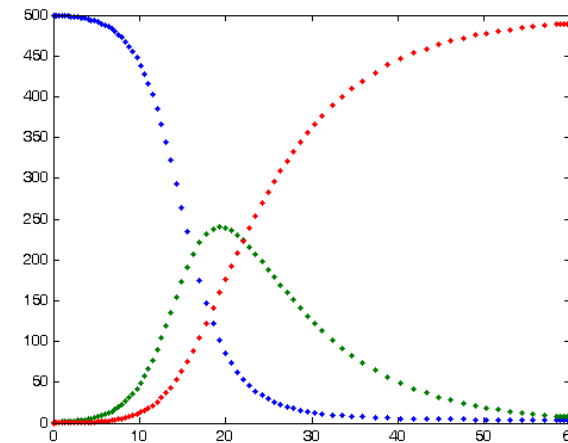
- SIR modell
- Természetesen tudni kell, hogy ki kivel érintkezik
- Legegyszerűbb a véletlen gráf

$$\frac{dS}{dt} = -\beta IS$$

$$\frac{dI}{dt} = \beta IS - \nu I$$

$$R_0 = \frac{\beta S}{\gamma}$$

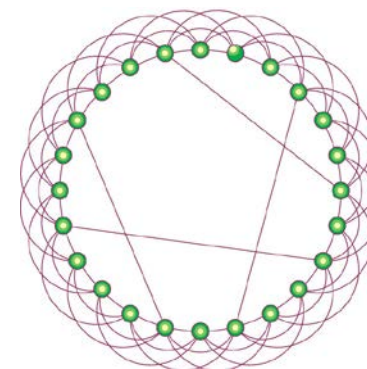
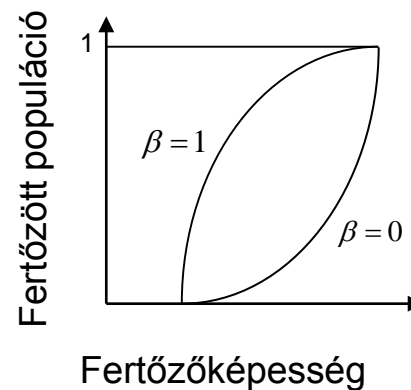
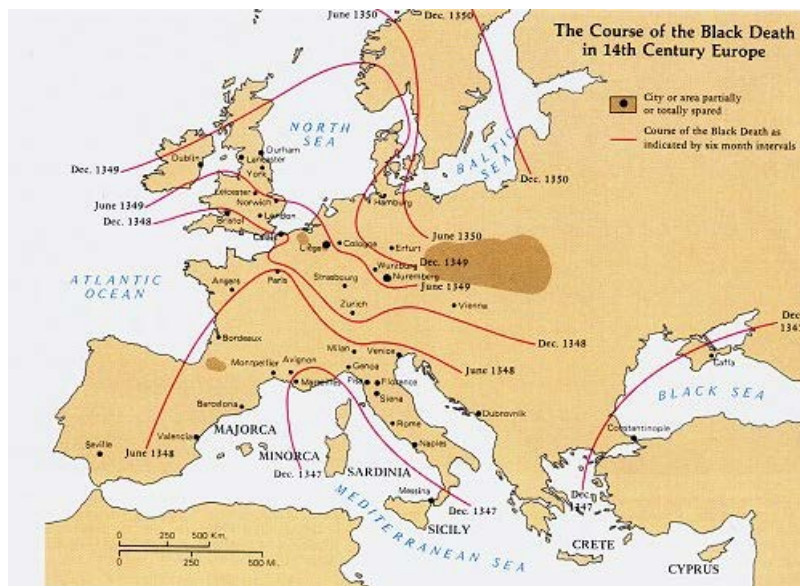
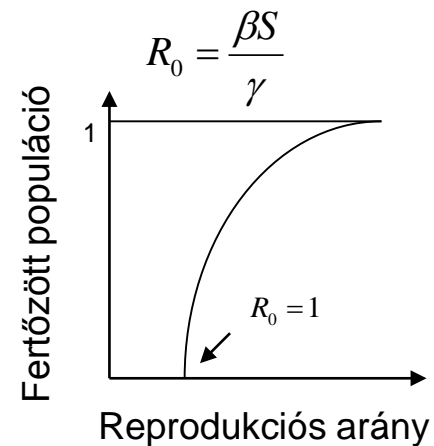
$$\frac{dR}{dt} = \nu I$$



Lassú, robbanás, lecsengés

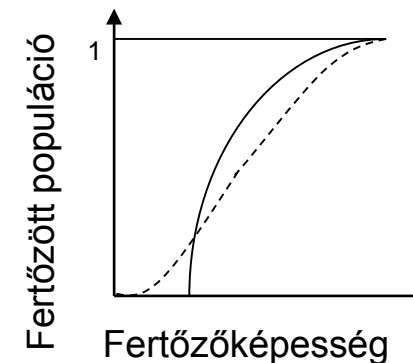
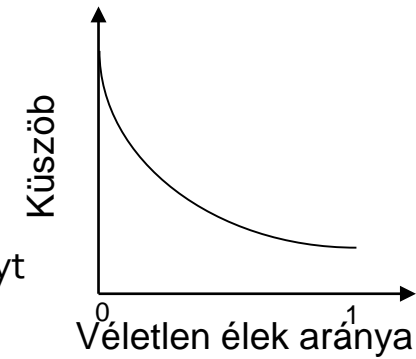
# Vírusterjedés modelleken

- Véletlen gráf esetén a reprodukciós arány teljesen meghatározza a lefolyást
  - Biztonságos szex
  - Állatok kivégzése (száj és körömfájás)
- Mi történik a W-S modellben
  - Véletlen gráf illetve rács esetén
  - Rács esetén csak az igazán durva betegség teljed el



# Vírusterjedés modelleken

- W-S modell esetén tehát
  - A shortcutokon keresztül gyorsan terjed a vírus
  - Új közösségeket megfertőzve (száj és körömfájás)
  - A kisvilágságot figyelmen kívül hagyva, az emberek nem érzik a veszélyt
  - Viszont van esély fellépni a kezdeti szakaszban
    - Modularitás mesterséges növelése
    - Reprodukciós arány csökkentése, immunizálás
  - Egy védekezési stratégia: a shortcutok elvágása
    - Tűcsere program
- Virus bulletin
  - A legtöbb számítógép vírus hosszan képes rejtőzködni a hálózatban
  - Hogy lehetséges ez? (SIR modellben nem lehet)
- Skálafüggetlen modell
  - Eltűnik a küszöb
  - Kegyetlen védekezési stratégia:
    - Hubok immunizálása
    - De hogy találjuk meg őket?



# Mit tehetünk még?

- Véletlen alany véletlen ismerősét immunizáljuk
- Számítógép vírusok
  - Microsoft minden kompatibilis mindennel
  - *„When you are dealing with rootkits and some advanced spyware programs, the only solution is to rebuild from scratch. In some cases, there really is no way to recover without nuking the systems from orbit”* Mike Danseglio, program manager in the Security Solutions group at Microsoft 2006
  - *"Detection is difficult, and remediation is often impossible,"* Danseglio declared. *"If it doesnt crash your system or cause your system to freeze, how do you know its there?"*
- Degenerált nem teljesen kompatibilis megoldások
- Heterogenitás



# Ajánlott szakirodalom

- **Evolution of Networks** – Dorogovtsev-Mendes
- **Statistical Mechanics of Complex Networks**  
Albert-Barabasi
- **The Structure and Function of Complex Networks**  
Mark Newman
- **És sok-sok cikk**

## További ajánlott irodalom

- Barabási Albert László – **Behálózva**
- Csermely Péter – **Rejtett hálózatok ereje**
- Duncan J. Watts - **Six degrees**