

Csoportosulások komplex hálózatokban

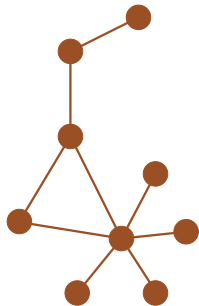
Palla Gergely

MTA-ELTE Statisztikus és Biológiai Fizika Kutatócsoport

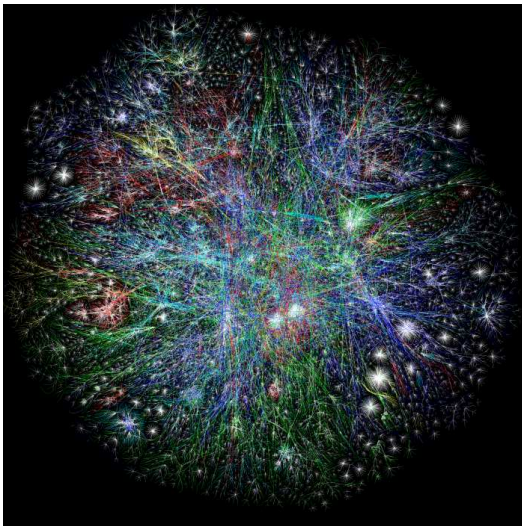
Az atomoktól a csillagokig

Bevezetés

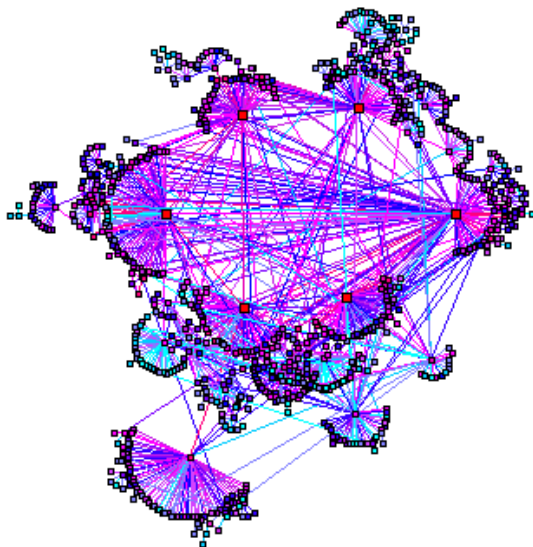
A hálózatos megközelítés



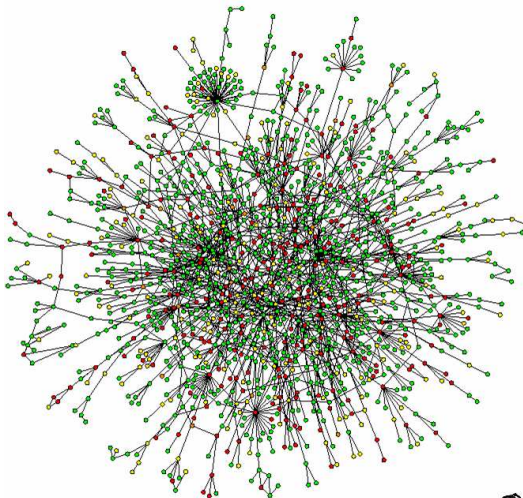
Példák valós hálózatokra



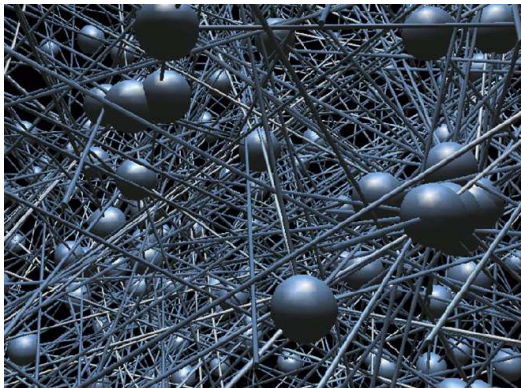
Példák valós hálózatokra



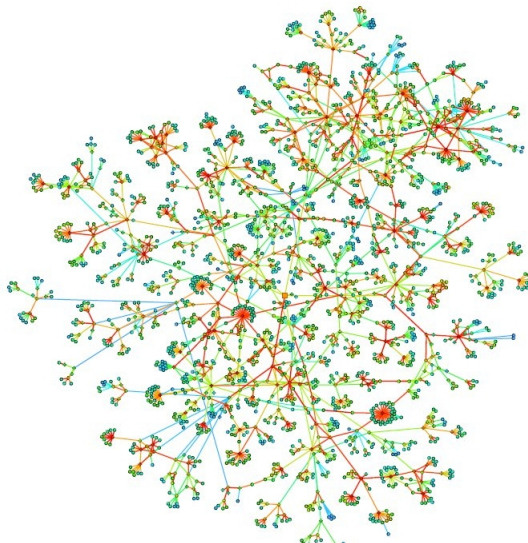
Példák valós hálózatokra



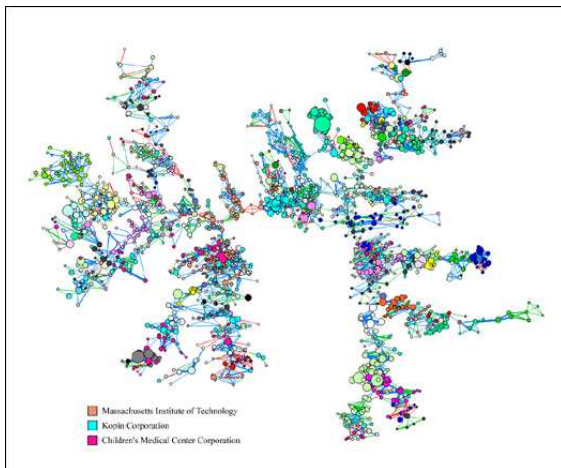
Példák valós hálózatokra



Példák valós hálózatokra



Példák valós hálózatokra



Példák valós hálózatokra



Mi dolga van egy fizikusnak a hálózatokkal?

A hálózat kutatás interdiszciplináris:

- matematika,
- biológia,
- szociológia,
- informatika,
- fizika,
 - statisztikus fizika
- (stb.)

Mi dolga van egy fizikusnak a hálózatokkal?

A hálózat kutatás interdiszciplináris:

- matematika,
- biológia,
- szociológia,
- informatika,
- fizika,
 - statisztikus fizika
- (stb.)

Mi dolga van egy fizikusnak a hálózatokkal?

A hálózat kutatás interdiszciplináris:

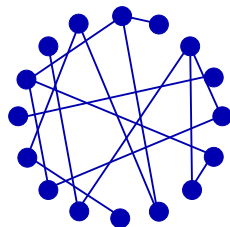
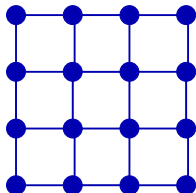
- matematika,
- biológia,
- szociológia,
- informatika,
- fizika,
 - **statisztikus fizika**
- (stb.)

Mi dolga van egy fizikusnak a hálózatokkal?

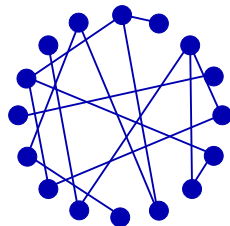
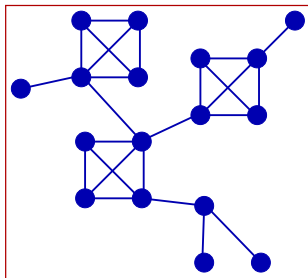
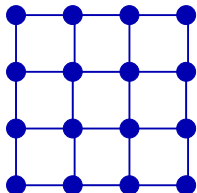
A hálózat kutatás interdiszciplináris:

- matematika,
- biológia,
- szociológia,
- informatika,
- fizika,
 - statisztikus fizika
- (stb.)

Hogyan képzeljük el egy valós hálózatot?



Hogyan képzeljük el egy valós hálózatot?



Hálózati csoportosulások

Csoportosulások, modulok, vagy klaszterek:
sűrűbben összekötött részek a hálózatban, melyeknek nincs egy általánosan elfogadott egyértelmű definíciója

Példák

- Egy család vagy egy baráti kör egy társas kapcsolat hálózatban,
- Azonos funkciójú, vagy azonos folyamatokban résztvevő fehérjék csoportja egy fehérje kölcsönhatási hálózatban,
- Hasonló témájú, egymáshoz sűrűn linkelt weblapok csoportja a világhálón

Hálózati csoportosulások

Csoportosulások, modulok, vagy klaszterek:
sűrűbben összekötött részek a hálózatban, melyeknek nincs egy általánosan elfogadott egyértelmű definíciója

Példák

- Egy család vagy egy baráti kör egy társas kapcsolat hálózatban,
- Azonos funkciójú, vagy azonos folyamatokban résztvevő fehérjék csoportja egy fehérje kölcsönhatási hálózatban,
- Hasonló témájú, egymáshoz sűrűn linkelt weblapok csoportja a világhálón

Hálózati csoportosulások

Csoportosulások, modulok, vagy klaszterek:

sűrűbben összekötött részek a hálózatban, melyeknek nincs egy általánosan elfogadott egyértelmű definíciója

Példák

- Egy család vagy egy baráti kör egy társas kapcsolat hálózatban,
- Azonos funkciójú, vagy azonos folyamatokban résztvevő fehérjék csoportja egy fehérje kölcsönhatási hálózatban,
- Hasonló témájú, egymáshoz sűrűn linkelt weblapok csoportja a világhálón

Hálózati csoportosulások

Csoportosulások, modulok, vagy klaszterek:

sűrűbben összekötött részek a hálózatban, melyeknek nincs egy általánosan elfogadott egyértelmű definíciója

Példák

- Egy család vagy egy baráti kör egy társas kapcsolat hálózatban,
- Azonos funkciójú, vagy azonos folyamatokban résztvevő fehérjék csoportja egy fehérje kölcsönhatási hálózatban,
- Hasonló témájú, egymáshoz sűrűn linkelt weblapok csoportja a világhálón

Hálózati csoportosulások

A csoportkeresés fontos, mert ...

- egy magasabb szerveződési szinten tanulmányozhatjuk a rendszer felépítését,
- segíthet tisztázni egyes ismeretlen funkciójú csúcsok rendszerben betöltött szerepét,
- pletykák és fertőzések gyorsabban terjednek a csoportokon belül mint a csoportok közt.

Hálózati csoportosulások

A csoportkeresés fontos, mert ...

- egy magasabb szerveződési szinten tanulmányozhatjuk a rendszer felépítését,
- segíthet tisztázni egyes ismeretlen funkciójú csúcsok rendszerben betöltött szerepét,
- pletykák és fertőzések gyorsabban terjednek a csoportokon belül mint a csoportok közt.

Hálózati csoportosulások

A csoportkeresés fontos, mert ...

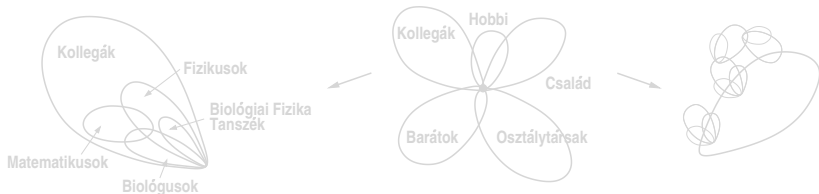
- egy magasabb szerveződési szinten tanulmányozhatjuk a rendszer felépítését,
- segíthet tisztázni egyes ismeretlen funkciójú csúcsok rendszerben betöltött szerepét,
- pletykák és fertőzések gyorsabban terjednek a csoportokon belül mint a csoportok közt.

Átfedő csoportosulások

A csoportok átfedhetnek egymással:

- Mindenki egyszerre tagja a családjának és a baráti körének,
- Egyes fehérijéknek több funkciója is van, így több csoportban is szerepelhetnek

A csoportok egymásba is ágyazódhatnak:

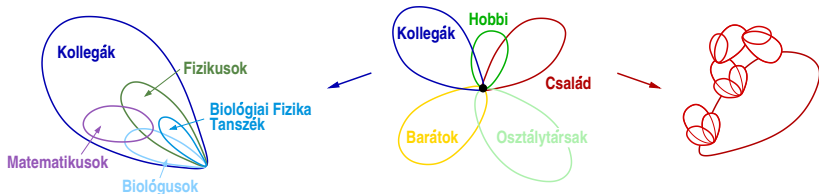


Átfedő csoportosulások

A csoportok átfedhetnek egymással:

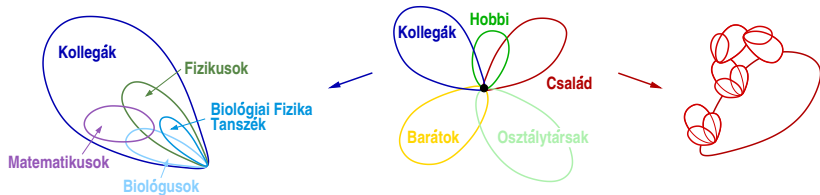
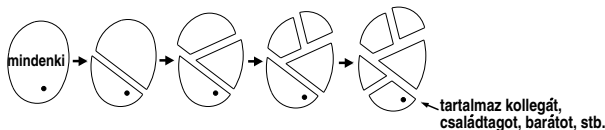
- Mindenki egyszerre tagja a családjának és a baráti körének,
- Egyes fehérijéknek több funkciója is van, így több csoportban is szerepelhetnek

A csoportok egymásba is ágyazódhatnak:



Átfedő csoportosulások

A csoportkereső módszerek nagyrésze izolált csoportokat keres:



A klikk perkolációs módszer

Az átfedéseket természetes módon engedi meg a **k -klikk perkoláción** alapuló csoport definíció.

Definíciók

- **k -klikk**: egy teljesen összekötött, k csúcsból álló részgráf.
- **k -klikk szomszédság**: két k -klikk szomszédos ha $k - 1$ csúcsuk, közös, (vagyis csak egy csúcsban térnek el).



A klikk perkolációs módszer

Az átfedéseket természetes módon engedi meg a **k -klikk perkoláción** alapuló csoport definíció.

Definíciók

- **k -klikk**: egy teljesen összekötött, k csúcsból álló részgráf.
- **k -klikk szomszédság**: két k -klikk szomszédos ha $k - 1$ csúcsuk, közös, (vagyis csak egy csúcsban térnek el).

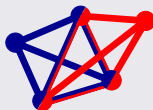


A klikk perkolációs módszer

Az átfedéseket természetes módon engedi meg a **k -klikk perkoláción** alapuló csoport definíció.

Definíciók

- **k -klikk**: egy teljesen összekötött, k csúcsból álló részgráf.
- **k -klikk szomszédság**: két k -klikk szomszédos ha $k - 1$ csúcsuk, közös, (vagyis csak egy csúcsban térnek el).



A klikk perkolációs módszer (CPM)

Definíciók

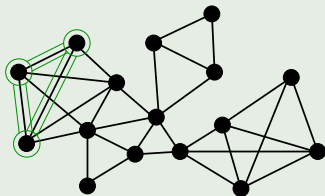
- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

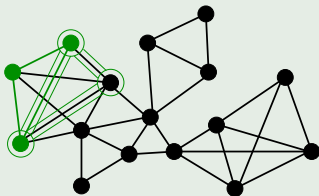


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás:** olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

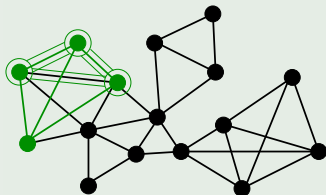


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

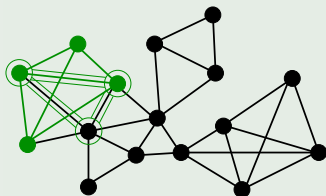


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

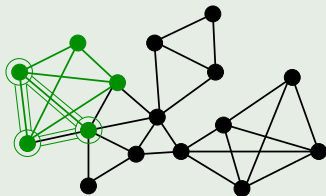


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

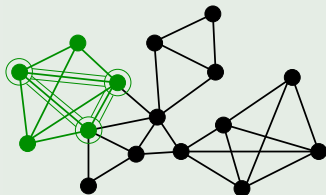


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

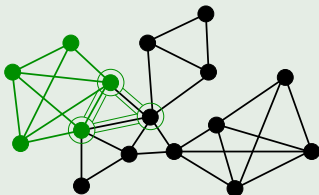


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

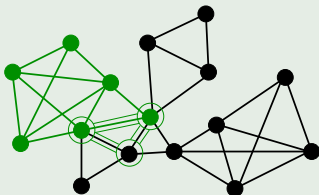


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

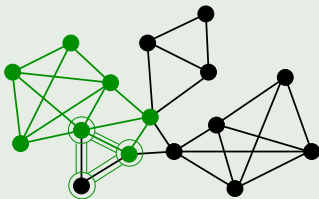


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

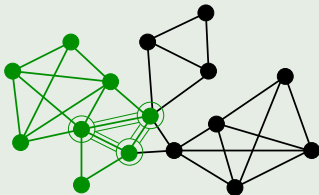


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

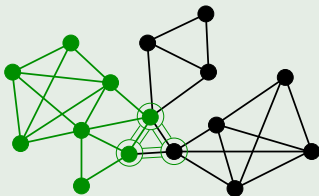


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás:** olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

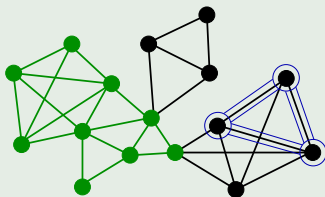


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

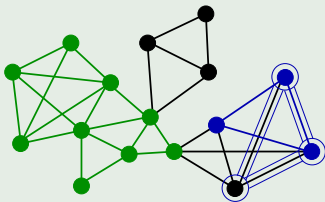


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

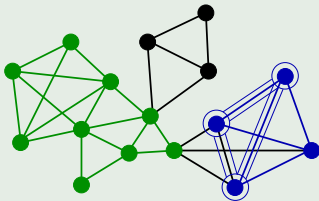


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

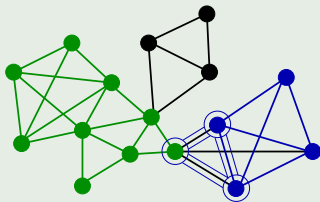


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

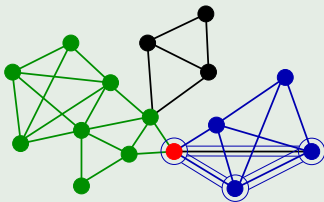


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

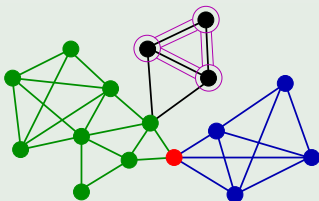


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

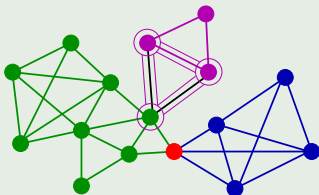


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

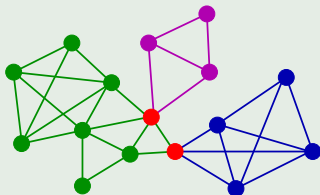


A klikk perkolációs módszer (CPM)

Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

Szemléltetés:

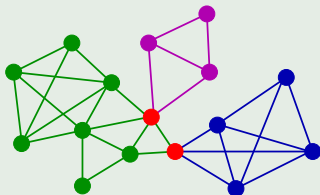


A klikk perkolációs módszer (CPM)

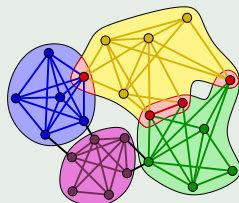
Definíciók

- **k -klikk csoportosulás**: olyan k -klikkek maximális uniója, melyben bármelyik k -klikkből eljuthatunk k -klikk szomszédságok sorozatán keresztül bármelyik másik k -klikkbe.

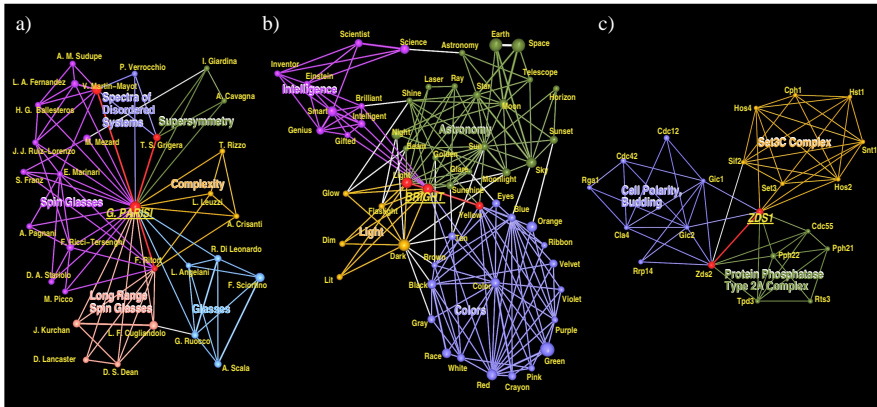
Szemléltetés:



ugyanaz $k = 4$ -re:

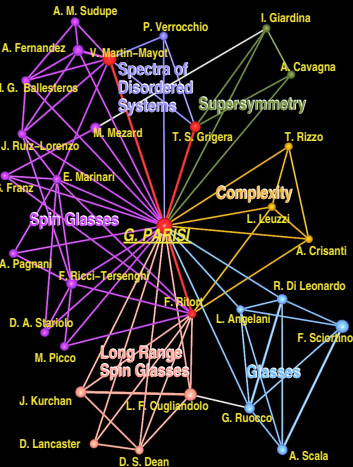


Egy adott csúcs csoportosulásai

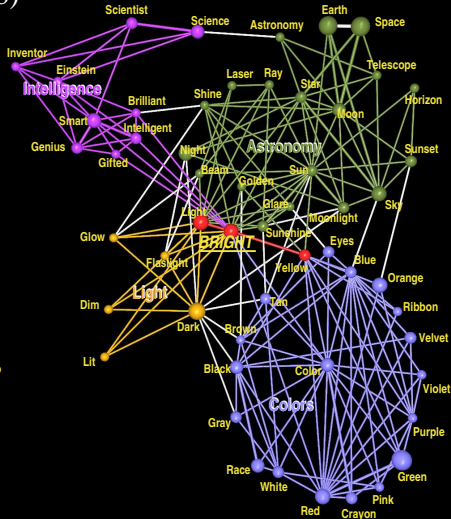


Egy adott csúcs csoportosulásai

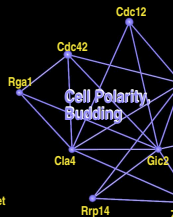
a)



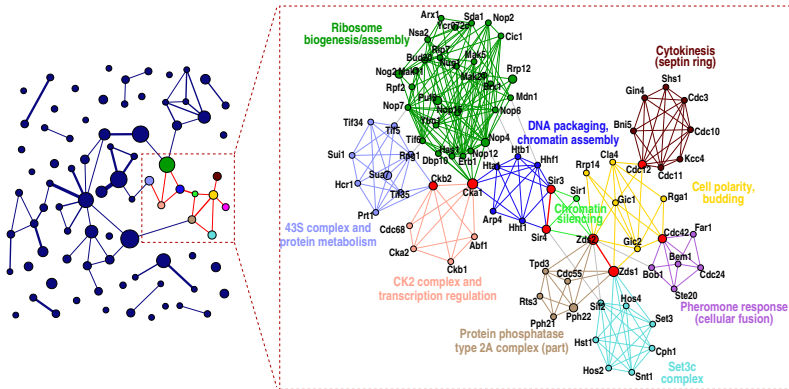
b)



c)

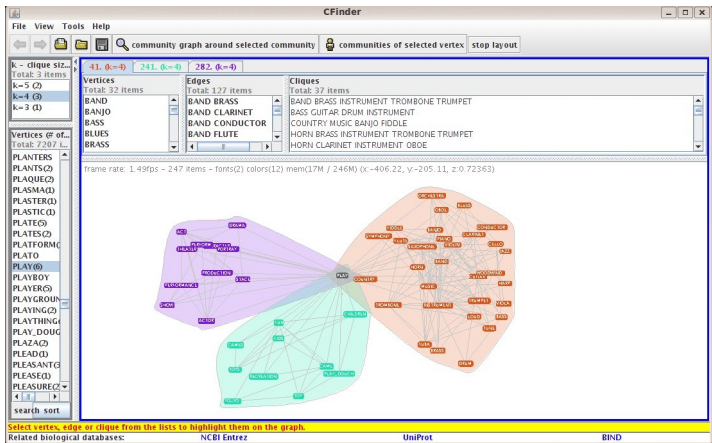


Csoportok hálózata



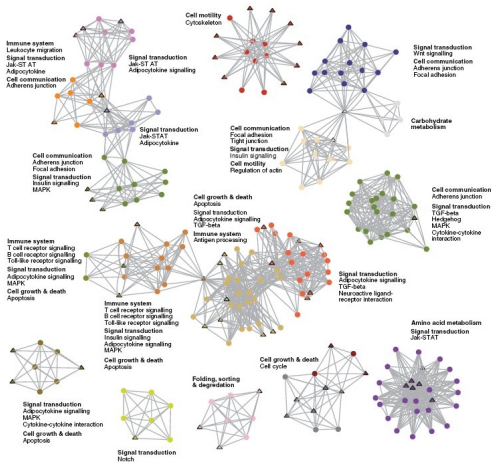
CFinder

A CPM-en alapuló hálózati csoportkereső és vizualizáló programcsomag: <http://www.cfinder.org>

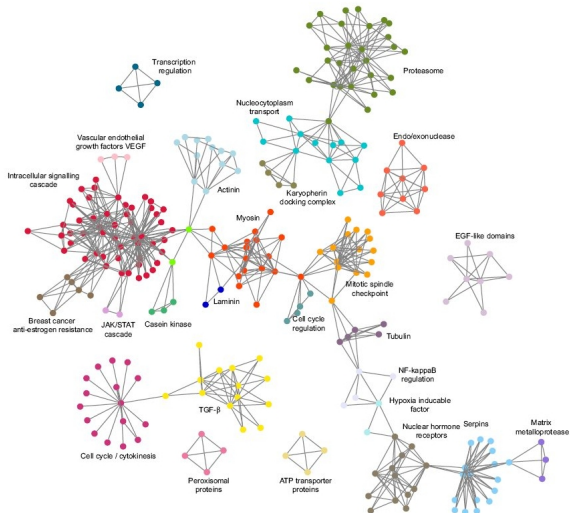


Alkalmazások:

Fehérje kölcsönhatási hálózatok, rákkutatás

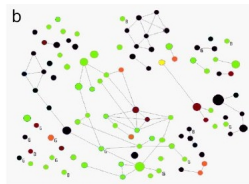
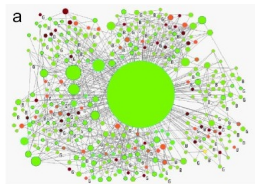


Alkalmazások: Fehérje kölcsönhatási hálózatok, rákkutatás



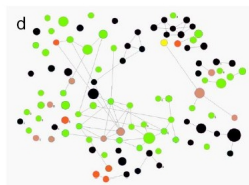
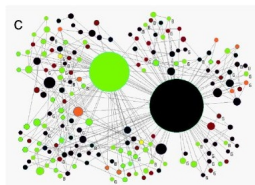
Alkalmazások:

Társas kapcsolat hálózatok

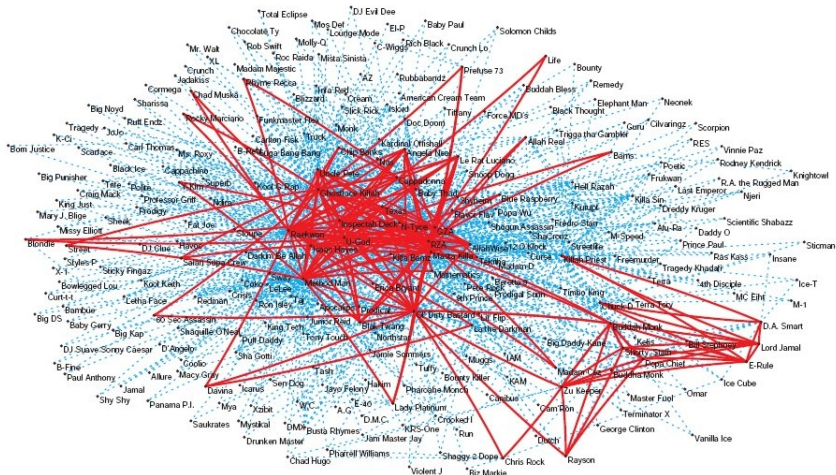


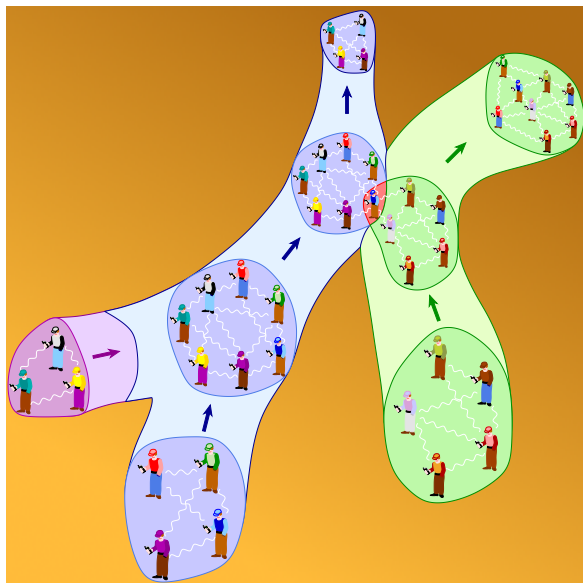
Node color

- Unknown
- Black
- Mixed
- Hispanic
- Asian
- White



Alkalmazások: Társas kapcsolat hálózatok





A tanulmányozott rendszerek

Az eredeti adatok:

- **A cond-mat cikk archívum, havi bontásban** (142 hónap, több mint 30000 szerző)
- **Egy mobiltelefon társaság (anonimizált) hívás listája** (52 hét, a hívásokat 2 hetes periódusokban aggregálták, több mint 4 millió felhasználó)

A tanulmányozott rendszerek

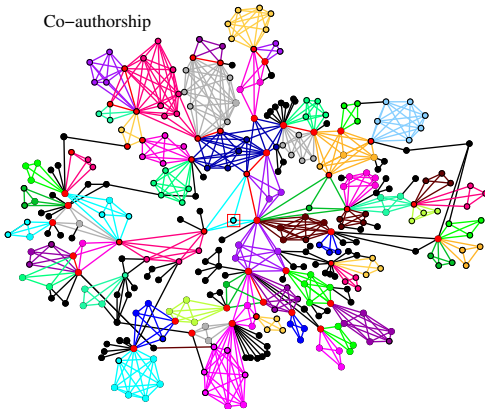
Az eredeti adatok:

- **A cond-mat cikk archívum, havi bontásban** (142 hónap, több mint 30000 szerző)
- **Egy mobiltelefon társaság (anonimizált) hívás listája** (52 hét, a hívásokat 2 hetes periódusokban aggregálták, több mint 4 millió felhasználó)

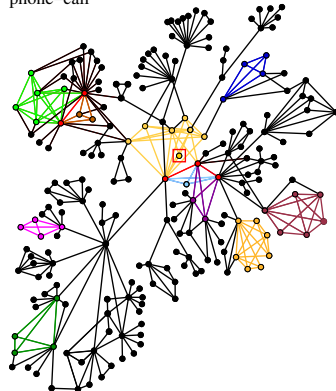
A lokális csoportosulás szerkezete

Egy véletlenszerűen választott csúcs kis környezetének statikus képe:

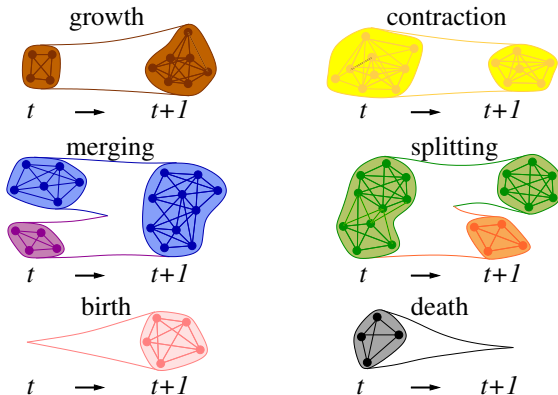
Co-authorship



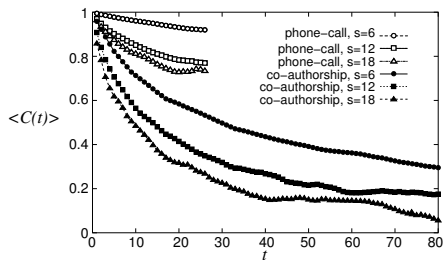
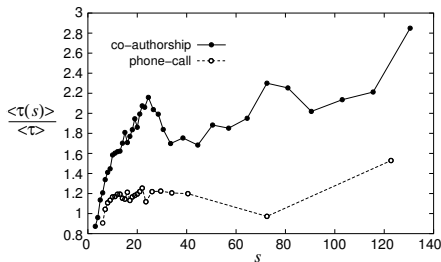
phone-call



A csoportfejlődés elemi eseményei



Auto-korrelációk



Egy adott A csoport **auto-korrelációja**:

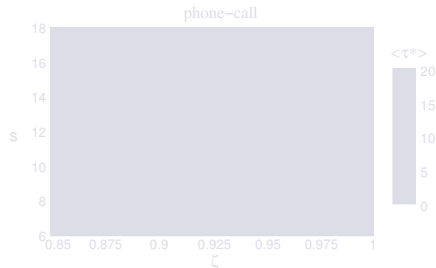
$$C_A(t) \equiv \frac{|A(t_0) \cap A(t_0 + t)|}{|A(t_0) \cup A(t_0 + t)|}$$

Stacionaritás és várható élettartam

Egy csoport **stacionaritása**:

$$\zeta \equiv \frac{\sum_{t=t_0}^{t_{\max}-1} C(t, t+1)}{t_{\max} - t_0 - 1}$$

A csoport **élettartama**, τ^* , tekinthető egy fajta „fitnessnek”.

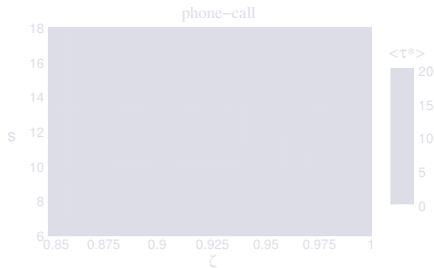


Stacionaritás és várható élettartam

Egy csoport **stacionaritása**:

$$\zeta \equiv \frac{\sum_{t=t_0}^{t_{\max}-1} C(t, t+1)}{t_{\max} - t_0 - 1}$$

A csoport **élettartama**, τ^* , tekinthető egy fajta „fitnessnek”.

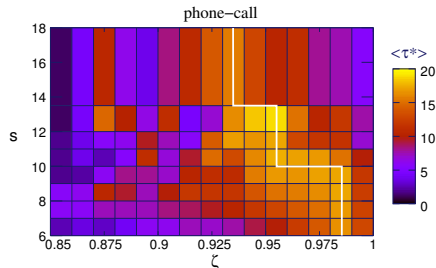
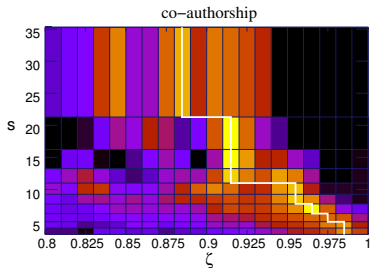


Stacionaritás és várható élettartam

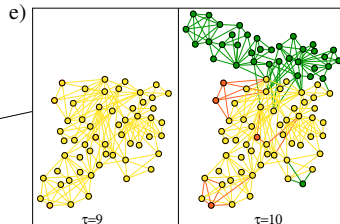
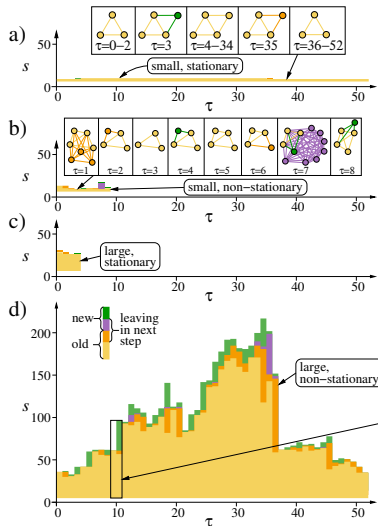
Egy csoport **stacionaritása**:

$$\zeta \equiv \frac{\sum_{t=t_0}^{t_{\max}-1} C(t, t+1)}{t_{\max} - t_0 - 1}$$

A csoport **élettartama**, τ^* , tekinthető egy fajta „fitnessnek”.



Csoportmérettől függő túlélési stratégia



Összefoglalás

- A hálózati csoportkeresés egy érdekes, izgalmas terület, széleskörű gyakorlati alkalmazásokkal
- A CPM egy effektív módszer átfedő hálózati csoportok felderítésére.
- A vizsgált időfejlődő hálózatokban eltérő túlélési stratégiát tapasztaltunk a kis- és a nagyméretű csoportok esetén.

Hálózat kutatási alkalmazások az "iparban"

- **Szervezetfejlesztési tanácsadás:**
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Hálózatkutatói alkalmazások az "iparban"

- Szervezetfejlesztési tanácsadás:
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Hálózat kutatási alkalmazások az "iparban"

- Szervezetfejlesztési tanácsadás:
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Hálózat kutatási alkalmazások az "iparban"

- Szervezetfejlesztési tanácsadás:
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Hálózat kutatási alkalmazások az "iparban"

- Szervezetfejlesztési tanácsadás:
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Hálózat kutatási alkalmazások az "iparban"

- Szervezetfejlesztési tanácsadás:
 - Kik a kulcsfigurák egy cég informális kapcsolathálózatában?
- Banki hálózatok:
 - Kik indíthatnak be egy lavina szerű csőd hullámot?
- Szolgáltatói hálózatok:
 - Kiken keresztül lehet a legjobban népszerűsíteni egy új terméket?

Köszönetnyilvánítás

- Vicsek Tamás



- Barabási
Albert-László



- Derényi Imre



- Farkas Illés



- Pollner Péter



- Ábel Dániel



OTKA K68669, K75334 és T049674,
NKTH, CellCom RET, Textrend,
Bolyai János Kutatási Ösztöndíj.

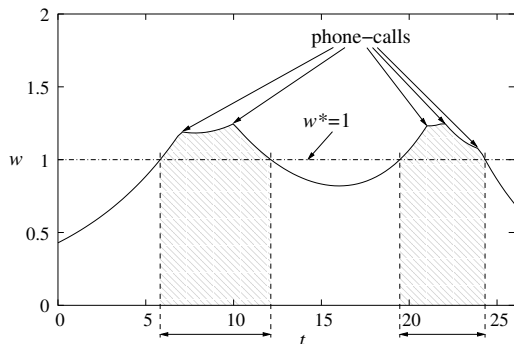
Időfejlődő kapcsolatok

A mobil telefonos adatoknál egy-két periódusból hiányoztak az adatok ...

Időfejlődő kapcsolatok

→ Időfejlődő élsúlyok:

$$w_{a,b}(t) = \sum_i w_i \exp(-\lambda |t - t_i| / w_i)$$



A csoportok nyomkövetése

Ha a csoportok csak növekednének...

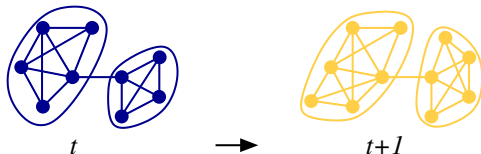
Egy pusztán növekedő hálózatban könnyű követni a CPM csoportosulásokat: egy t -ben megtalált csoportot biztosan tartalmazni fog egy és csak is egy $t + 1$ -ben megtalált csoport.



A csoportok nyomkövetése

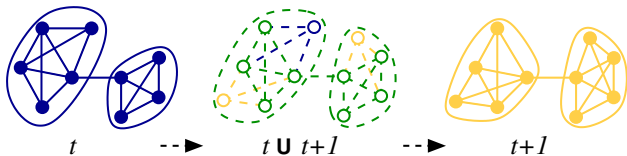
Ha a csoportok csak növekednének...

Egy pusztán növekedő hálózatban könnyű követni a CPM csoportosulásokat: egy t -ben megtalált csoportot biztosan tartalmazni fog egy és csak is egy $t + 1$ -ben megtalált csoport.



A csoportok nyomkövetése

A módszer



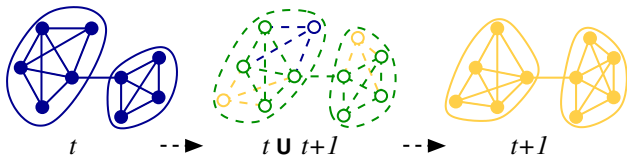
A szomszédos időlépéseknél talált csoportokat
összeparosíthatjuk

- a két időlépésnél látott hálózat uniójában talált csoportok alapján.
- Ha több párosítás is felmerül, akkor a csoportok relatív átfedése alapján döntünk.

$$C(A, B) \equiv \frac{|A \cap B|}{|A \cup B|}$$

A csoportok nyomkövetése

A módszer



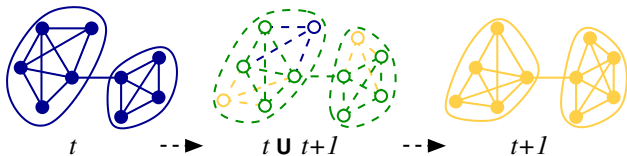
A szomszédos időlépéseknél talált csoportokat összepárosíthatjuk

- a két időlépésnél látott hálózat uniójában talált csoportok alapján.
- Ha több párosítás is felmerül, akkor a csoportok relatív átfedése alapján döntünk.

$$C(A, B) \equiv \frac{|A \cap B|}{|A \cup B|}$$

A csoportok nyomkövetése

A módszer



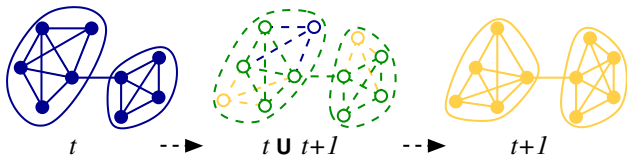
A szomszédos időlépéseknél talált csoportokat
összefűzhetjük

- a két időlépésnél látott hálózat uniójában talált csoportok alapján.
- Ha több párosítás is felmerül, akkor a csoportok relatív átfedése alapján döntünk.

$$C(A, B) \equiv \frac{|A \cap B|}{|A \cup B|}$$

A csoportok nyomkövetése

A módszer

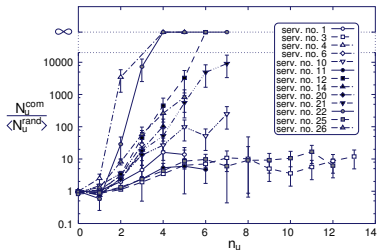
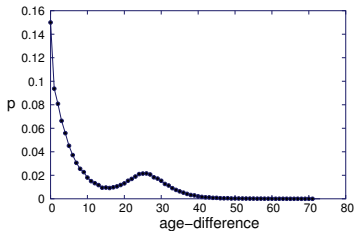
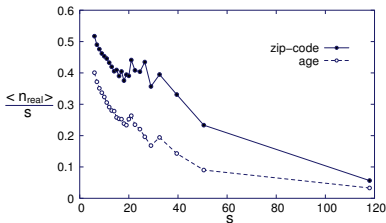
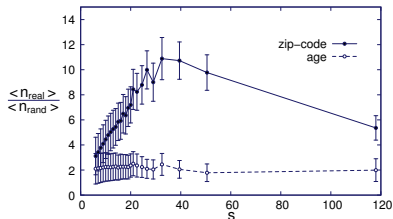


A szomszédos időlépéseknél talált csoportokat összepárosíthatjuk

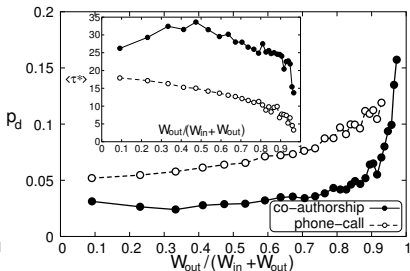
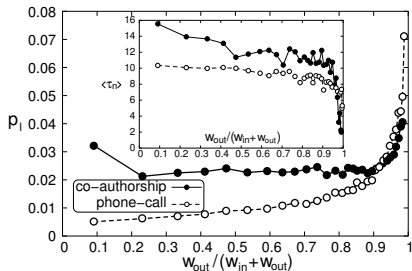
- a két időlépésnél látott hálózat uniójában talált csoportok alapján.
- Ha több párosítás is felmerül, akkor a csoportok relatív átfedése alapján döntünk.

$$C(A, B) \equiv \frac{|A \cap B|}{|A \cup B|}$$

Csoport ellenőrzés

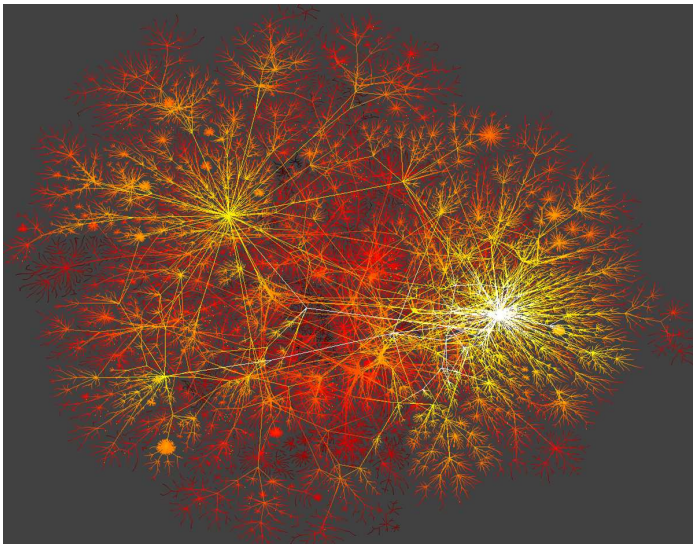


Elkötelezettség és stabilitás

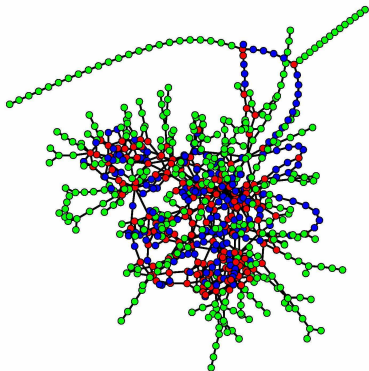


- W_{in} : egy tagnak a saját csoportja felé menő összesített élsúlya,
- W_{out} : egy tagnak a hálózat egyéb részei felé menő összesített élsúlya,
- W_{in} : egy csoporton belüli élek összesített súlya,
- W_{out} : egy csoportból a csoporton kívülre vezető élek összesített súlya.

Példák valós hálózatokra



Példák valós hálózatokra



Példák valós hálózatokra

