TÁMOP-4.1.1.F-14/1/KONV-2015-0006

TÁRSADALMI ÉS GAZDASÁGI HÁLÓZATOK MODELLEZÉSE

6. ELŐADÁS: HÁLÓZATOK NÖVEKEDÉSI MODELLJEI:

‚UNIFORM ÉS PREFERENTIAL ATTACHMENT’

London András

SZÉCHENYI 2020
MAGYARORSZÁG KORMÁNYA
Európai Unió
Európai Szociális Alap
BEFEKTETÉS A JÖVŐBE

Hogyan nőnek a hálózatok?

- Statikus hálózatos modellek: a pontok száma (n) fix, az élket ‘valamilyen véletlen’ generálja
  - Erdős-Rényi modell (kis távolságok, alacsony klaszterezettég)
  - Watts-Strogatz modell (kis távolságok, magas klaszterezettség)
  - Konfigurácó modell (adott fokszámsorozatú gráf)
  - Sztochasztikus Blokk Model (adott magasszintű struktúra)
- Ugyanakkor sokszor valós dinamikus rendszereket modellezünk hálózattal
  - gondoljuk a web gráf növekedésére
  - a baráti és munkahelyi kapcsolatok kialakulására
  - a tudományos publikációk citációira

Komplex hálózatok szerkezetét vizsgálva (fokszámeloszlás, közösségszerkezet, centralitások, stb.) számos tulajdonságot megtudhatunk a modellezett rendszerről, **DE**

nem feltétlenül tudjuk, hogy miért pont ezek a mintázatok jelennek meg?

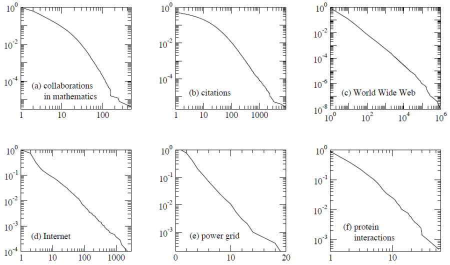
- Szociális hálókban miért nagy a klaszterezettség?
- Biológiai rendszerekben miért jelentős a mag-periféria szerkezet?
- Tudományos publikációk citációs hálózata miért követ hatványtörvényes fokszámeloszlást
- Online közösségi hálókban miért jelennek meg 'vastag-farkú' fokszámeloszlások?

$\Longrightarrow$ Milyen mechanizmus hozta létre ezeket a hálózatokat?

Hatványtörvény

$$\mathbb{P}(k_i = k) = ck^{-\alpha}$$

- Városok lakossága
- Szexuális partnerek száma
- Gének kópiáinak száma egy genomban
- stb.



A hatványtörvények története

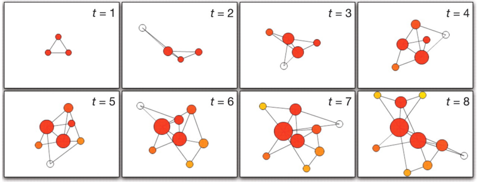
- Pareto, 1897: Pareto eloszlás („80-20” törvény): a javak 80%-át a lakosság 20%-a birtokolja
- Zipf, 1916: szavak gyakorisága szövegekben, városok lakossága (a j-edik leggyakoribb angol szó gyakorisága az összes szövegben $1/j$-vel arányos)
- Simon, 1955: „a gazdag még gazagabb lesz” (the rich gets richer)
- Price, 1965: citációs hálózatok vizsgálata; az ötlet: egy tudományos cikk annál több idézést kap, minél több idézést kapott már eddig → „kumulatív előny”
- Albert Réka és Barabási László, 1999: **preferential attachment**

Barabási-Albert modell [1]

Preferential attachment dinamikus modell:

1. kezetben egy összefüggő G_0 gráf n_0 ponton
2. t időpontban hozzáadunk G_t-hez egy új v pontot úgy, és m_0 élt v-ből G_{t-1}-be, hogy

$$\mathbb{P}(v\text{-t összekötjük egy meglévő } i\text{-vel}) = \frac{k_i}{\sum_j k_j}$$



[1] Barabási & Albert, *Science*, 1999

Barabási-Albert modell

Ebből

$$\mathbb{P}(\text{létező } i \text{ pont „kap" új élt } t \text{ időpontban}) = m_0 \frac{k_i}{\sum_j k_j}$$

t-ben összesen tm_0 él van a gráfban $\Longrightarrow \sum_{j=1}^{t} k_j(t) = 2tm_0$. A kettőből adódik, hogy annak a valószínűsége, hogy az i pont kap új élt t-ben ($t = 1, 2, \dots$): $\frac{k_i(t)}{2t}$

- Kis „csalással" a várható fokszám időbeli változását a

$$\frac{dk_i(t)}{dt} = \frac{k_i(t)}{2t}$$

differenciálegyenlet írja le a $k_i(i) = m_0$ (az i-edik pontot i időpillanatban adtuk hozzá a gráfhoz) kezdeti feltétellel (feltéve, hogy a fokszám folytonos valószínűségi változó $\leftarrow$ ez csalás!)

Barabási-Albert modell

Az egyenlet megoldsása:

$$k_i(t) = m_0 \left(\frac{t}{i}\right)^{1/2}$$

A fokszámeloszlás meghátározásához meg kellene nézni, hogy t-ben hány pont foka kisebb vagy egyenlő k-val:

$$\mathbb{P}(k_i(t) < k) = \mathbb{P}(m_0 \left(\frac{t}{i}\right)^{1/2} < k)$$

Ebből

$$\mathbb{P}(i > m_0^2 k/t^2) = 1 - \mathbb{P}(i \leq m_0^2 t/k^2) = 1 - m_0^2 t/k^2(t + m_0)$$

feltéve, hogy a pontokat egyenlő időintervallumokon adjuk a rendszerhez

Barabási-Albert modell

A sűrűségfüggvény ebből

$$\mathbb{P}(k) = \frac{dP(k_i(t) < k)}{dk}$$

mely stacionárius megoldása

$$\mathbb{P}(k) = 2\frac{m_0^2}{k^3} \sim k^{-3}$$

Azaz a modell egy skálafüggetlen fokszámeloszlású hálózatot generál.

Uniform attachment

- az i cimkéjű pont $t = i$ időpontban születik $(i = 1, 2, \dots)$
- $k_i(t)$ az i pont foka t-ben
- kezdetben m_0 pont
- $k_i(i) = m_0$ kezdeti feltétel Minden t időpillanatban az újonnan születő pont m_0 új éllel kötődik a már meglévő t ponthoz véletlen szerűen, ezért $t > i$-ben az i pont várható fokszáma:

$$\frac{dk_i(t)}{dt} = \frac{m}{t}$$

Ezután ugyanaz a sztori, mint az előbb. (szorgalmi feladat)

Vertex copy

- Adott egy G_0 hálózat
- Válasszunk egy pontot ki véletlen szerűen, „másoljuk le” az összes élével együtt
- Minden élre dobjunk fel egy érmét: ha fej (q-val), akkor ugyanahhoz a ponthoz kössük be, ahova az eredeti pont esetén volt kötve, ha írás $(1 - q)$, akkor véletlenül választott ponthoz kössük be
  $\Longrightarrow$ Itt is a hatványtörvényes fokszámeloszlás jön elő (projekt feladat)