Distances in preferential attachment graphs

Sander Dommers

Joint work with:
Remco van der Hofstad
Gerard Hooghiemstra

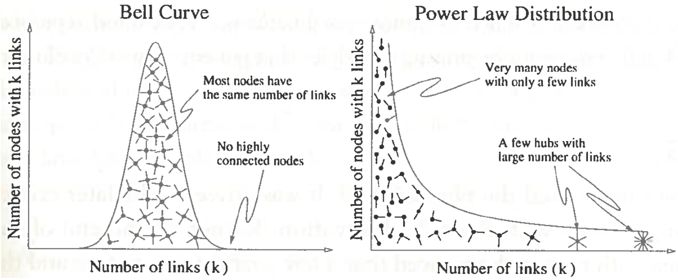
Technische Universiteit
Eindhoven
University of Technology

Fundamentals of Future Networking
November 8, 2011

Properties of complex networks

Power-law behavior

Number of vertices with degree k is proportional to $k^{-\tau}$.



Barabási, Linked, '02

Small worlds

Distances in the network are small

Preferential attachment

Possible explanation (Barabási, Albert, *Science*, 1999):

- "Networks *expand* continuously by the addition of new vertices;
- New vertices attach *preferentially* to sites that are already *well connected.*"

Linear preferential attachment

Here, *linear* preferential attachment:

Probability of connecting to certain vertex
proportional to its *degree* plus a *constant* δ.

Gives a *power-law* degree sequence (Cooper, Frieze, 2003):

Number of vertices with degree k is proportional to $k^{3+\delta/m}$,

where m is the number of edges of a new vertex.

Technische Universiteit Eindhoven University of Technology

Model definition

Algorithm to construct the preferential attachment graph:

- Fix the number of edges per new vertex $m \geq 1$ and constant $\delta > -m$.
- At time $t = 2$, start with 2 vertices connected by $2m$ edges.
- At time $t + 1$, given the graph at time t, $\mathrm{PA}(t)$, add a vertex and let, for $1 \leq j \leq m$,

$$\mathbb{P}[j\text{th edge of } (t+1) \text{ is connected to } i | \mathrm{PA}(t)] = \frac{D_i(t) + \delta}{(2m + \delta)t}.$$

Known results

Let

$$\mathrm{diam}_t(G(V,E)) = \max_{i,j\in V} \mathrm{dist}_t(i,j).$$

Then

- For $m = 1$, the graph is a tree and (Pittel, 1994):

$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \log t;$$

- For $m \geq 2, \delta = 0$ (Bollobás, Riordan, 2004):

$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \frac{\log t}{\log\log t}.$$

TU/e Technische Universiteit Eindhoven University of Technology

Results

Theorem (DvdHH, 2010)

- *For* $m \geq 2, \delta > 0$,

$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \log t;$$

- *For* $m \geq 2, \delta < 0$,

$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \log \log t.$$

Results

Theorem (DvdHH, 2010)

- *For* $m \geq 2, \delta > 0$,

$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \log t;$$

- *For* $m \geq 2, \delta < 0$,

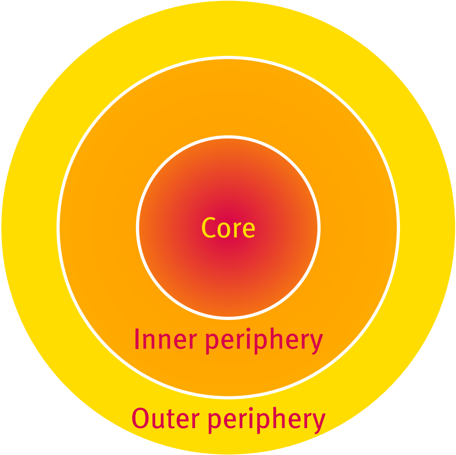
$$\mathrm{diam}_t(\mathrm{PA}(t)) \sim \log \log t.$$

Sketch of proof that, for $m \geq 2, \delta < 0$,

$$\mathrm{diam}_{2t}(\mathrm{PA}(2t)) \leq \left(\frac{4}{|\log(\tau - 2)|} + \frac{4\sigma}{m} \right) \log \log t.$$

Technische Universiteit Eindhoven University of Technology

Split in layers

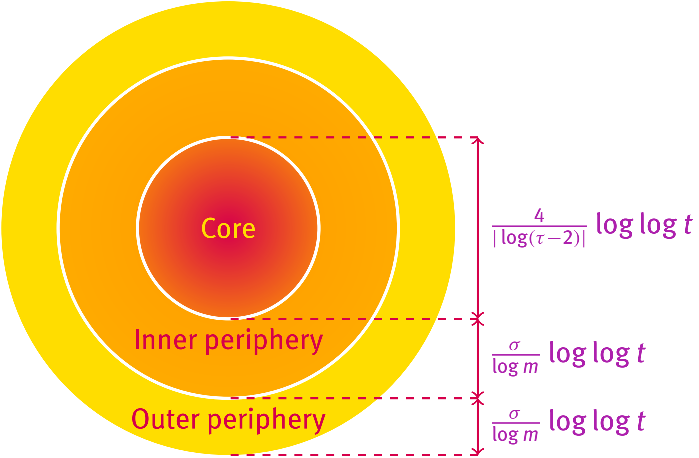


Core: vertices with degree at least $(\log t)^{\sigma}$, for $\sigma > \frac{1}{3-\tau}$;

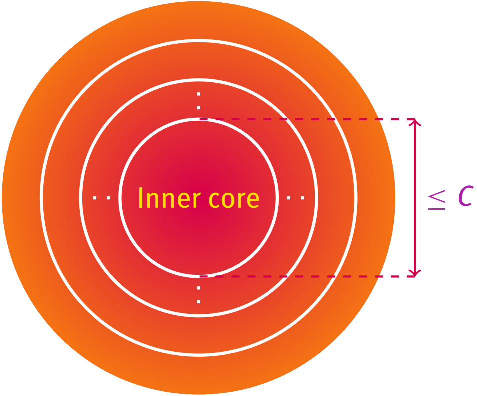
Inner periphery: vertices in $\{1, \ldots, t\} \backslash \text{Core}$;

Outer periphery: vertices in $\{t+1, \ldots, 2t\}$.

Distances in/between layers



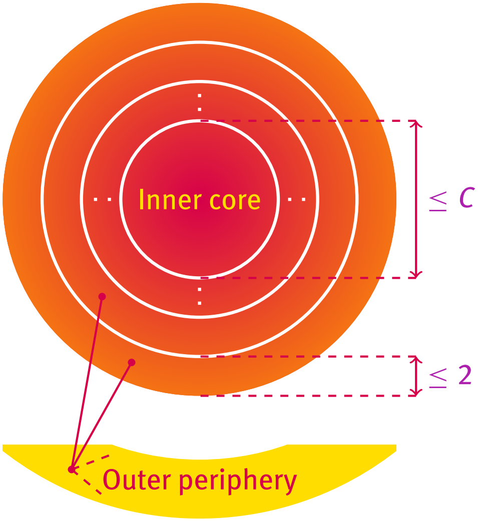
Diameter of the core



Split core in layers of *decreasing degree.*

Number of layers:
$\log\log t \big/ |\log(\tau-2)|$.

Diameter of the core



Split core in layers of *decreasing degree.*

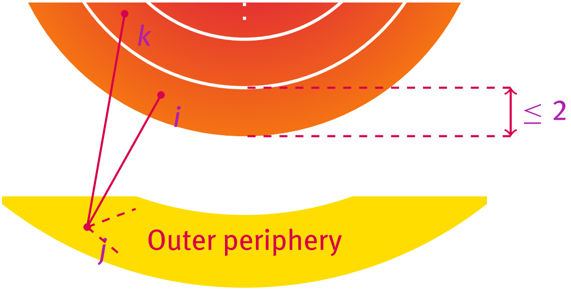
Number of layers:
$\log\log t \Big/ |\log(\tau-2)|.$

Connection between layers via a *t-connector.*

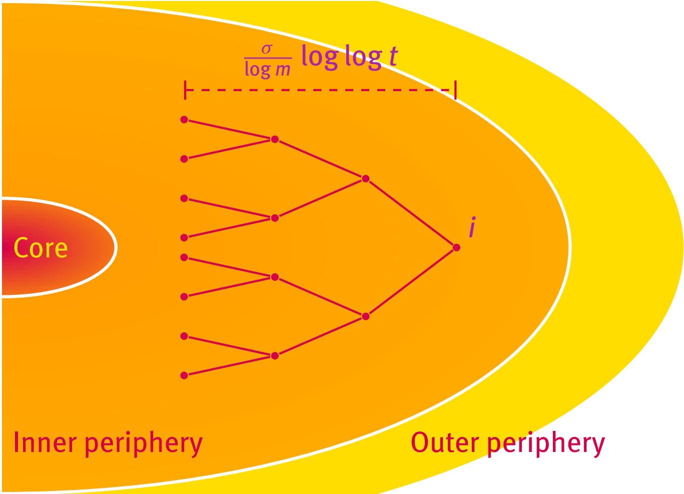
t-connectors

Vertex j is a *t-connector* between vertex i and a set of vertices A if

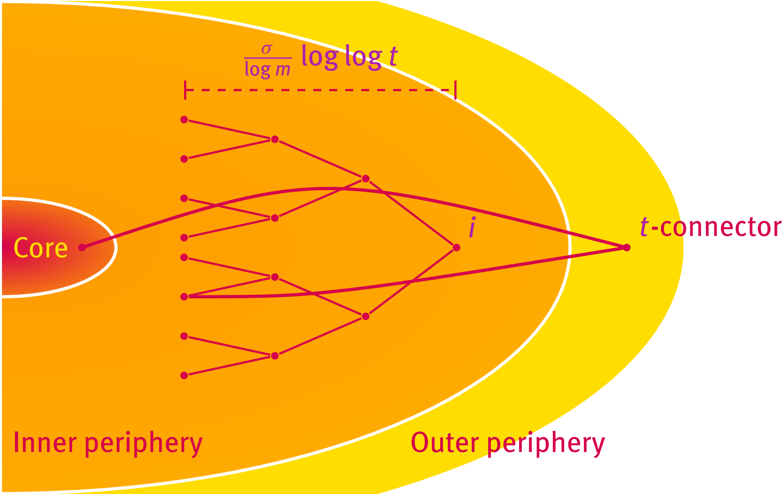
- the *first* edge of j connects to i,
- the *second* edge of j connects to k, for *some* $k \in A$.



Distance from inner periphery to core

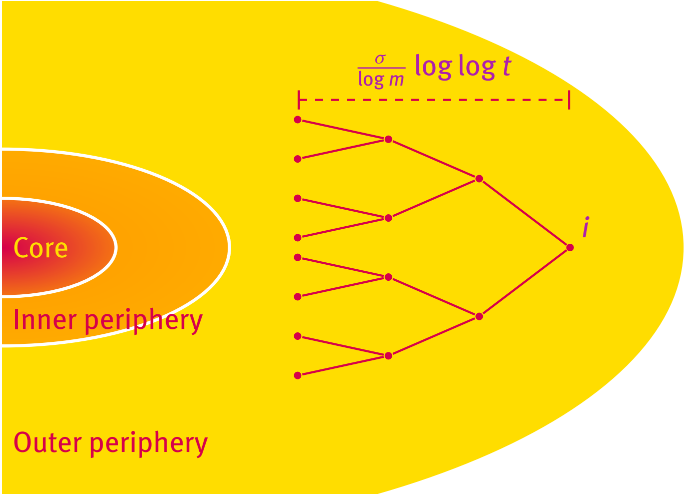


Distance from inner periphery to core



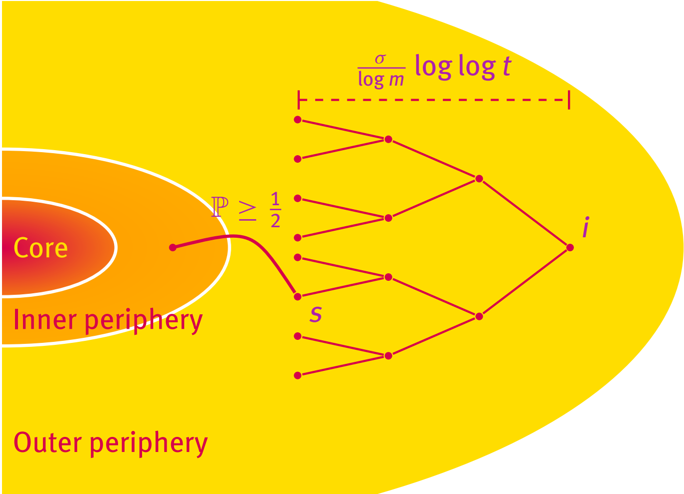
Technische Universiteit Eindhoven University of Technology

Distance from outer to inner periphery

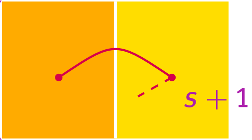


TU/e Technische Universiteit Eindhoven University of Technology

Distance from outer to inner periphery



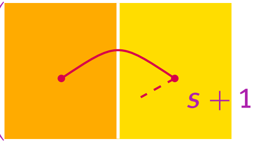
Probability of connecting to the inner periphery



$$\mathbb{P}\left(\qquad\right) = \sum_{i=1}^{t} \frac{D_i(s)+\delta}{(2m+\delta)s}$$

Probability of connecting to the inner periphery

$\mathbb{P}\Big($



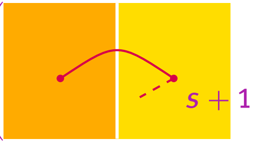
$\Big)$

$$= \sum_{i=1}^{t} \frac{D_i(s)+\delta}{(2m+\delta)s}$$

$$\geq \sum_{i=1}^{t} \frac{D_i(t)+\delta}{(2m+\delta)2t} = \frac{2mt+\delta t}{(2m+\delta)t} = \frac{1}{2}.$$

Probability of connecting to the inner periphery

$$\mathbb{P}\left(\quad \right)$$



$$= \sum_{i=1}^{t} \frac{D_i(s)+\delta}{(2m+\delta)s}$$

$$\geq \sum_{i=1}^{t} \frac{D_i(t)+\delta}{(2m+\delta)2t} = \frac{2mt+\delta t}{(2m+\delta)t} = \frac{1}{2}.$$

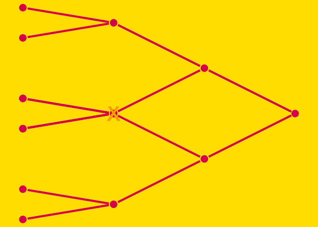
Number of leafs equals

$$m^{\frac{\sigma}{\log m}\log\log t} = (\log t)^{\sigma}.$$

Since $\sigma = \frac{1}{3-\tau} > 1$, probability of no connection between exploration tree and inner periphery is $o\left(\frac{1}{t}\right)$.

Collisions in the exploration tree

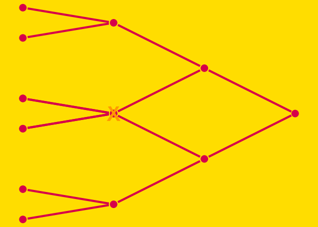
Possibility of collisions in the exploration tree:



$$\mathbb{P}[\text{at least 1 collission}] \leq \sum_{s \in \mathcal{T}} \frac{D_s(2t) + \delta}{t(2m + \delta)}$$

Collisions in the exploration tree

Possibility of collisions in the exploration tree:

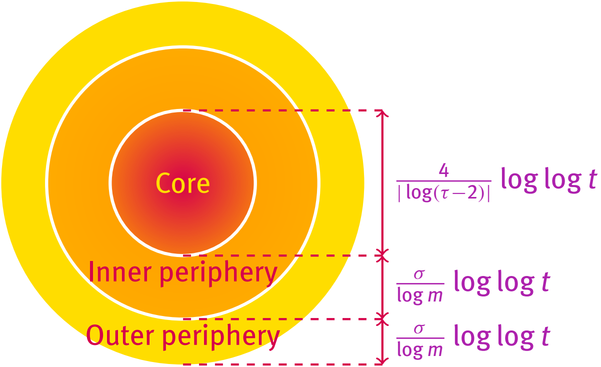


$$\mathbb{P}[\text{at least 1 collission}] \leq \sum_{s \in \mathcal{T}} \frac{D_s(2t) + \delta}{t(2m + \delta)} \leq m^{|\mathcal{T}|} \frac{(\log t)^{\sigma}}{t} \leq \frac{m(\log t)^{2\sigma}}{t}.$$

Multiple collisions are very rare!

Technische Universiteit Eindhoven University of Technology

Summary



Indeed,

$$\mathrm{diam}_{2t}(\mathrm{PA}(2t)) \leq \left(\frac{4}{|\log(\tau-2)|} + \frac{4\sigma}{m} \right) \log\log t.$$

TU/e Technische Universiteit Eindhoven University of Technology