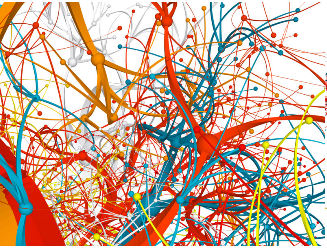
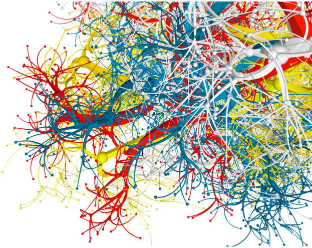
# Modeling Community Growth: densifying graphs or sparsifying subgraphs?

Róbert Pálovics

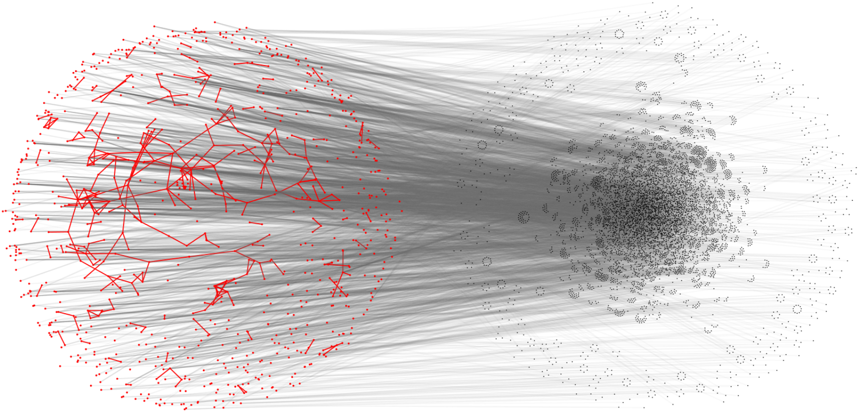




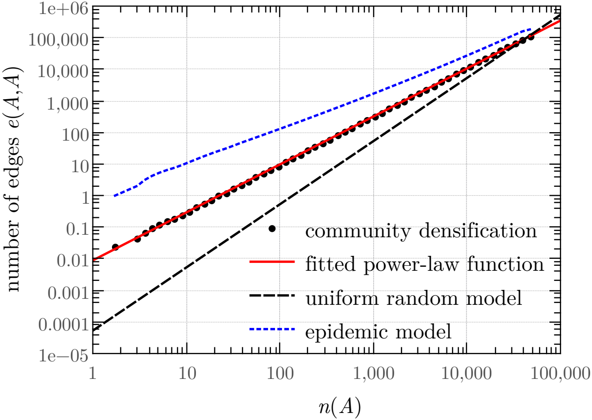
May 12, 2014

# Outline

# Subgraph Densification laws

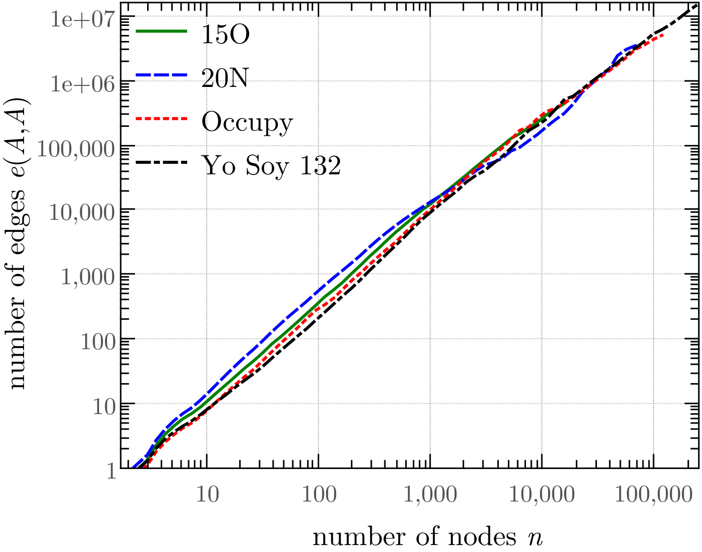


# Subgraph Densification laws




$$\bar{d} = \frac{e}{n} \qquad \rho = \frac{2e}{n(n-1)} \sim \frac{e}{n^2}$$

# Subgraph Densification laws



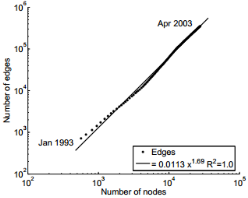
# Empirical observations in Large Networks

| | Unweighted |
|---|---|
| **Static** | **Power-law degree distribution(Barabasi)**<br>Triangle Power Law (Tsourakakis)<br>Eigenvalue Power Law (Siganos)<br>Community structure (Girvan, Newman, Flake) |
| **Dynamic** | **Densification Power Law (Falutsos, Leskovec, Dorogovtsev)**<br>Small and Shrinking Diameters (Barabasi, Leskovec)<br>Constant Size 2nd and 3rd connected components (McGlohon)<br>Principal Eigenvalue Power Law (Akoglu)<br>Burstiness |

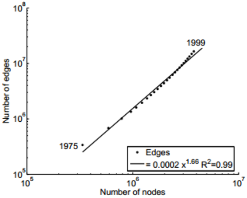
| | Weighted |
|---|---|
| **Static** | Snapshot Power Law (McGlohon) |
| **Dynamic** | Weight Power Law (McGlohon) |

# Densification laws (Kleinberg, Leskovec, Faloutsos)

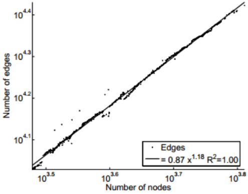
*Densification Power Law:* the number of nodes $N$ and the number of edges $E$ should follow a power-law in the form of $E(t) \propto N(t)^{\gamma}$, with $\gamma > 1$, over time.

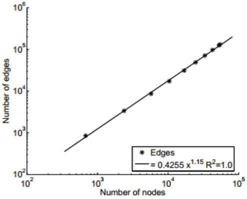



(a) arXiv




(b) Patents

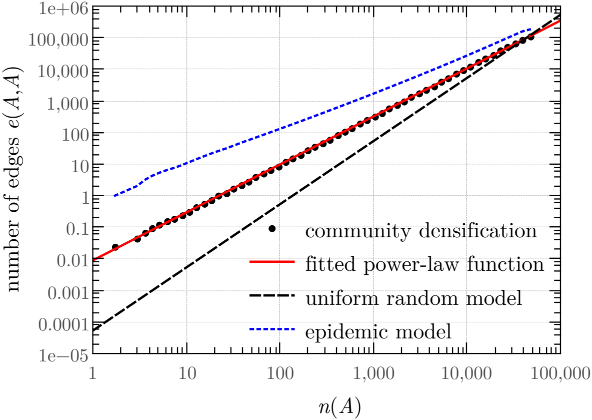


# Densification laws (Kleinberg, Leskovec, Faloutsos)

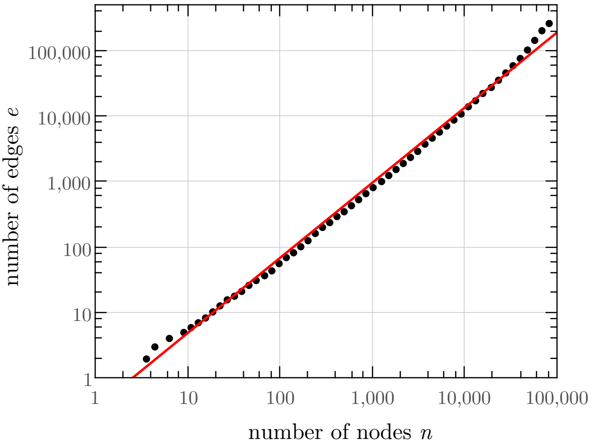
- Graphs over Time: Densication Laws, Shrinking Diameters and Possible Explanations
- *Densification Power Law:* the number of nodes $N$ and the number of edges $E$ should follow a power-law in the form of $E(t) \propto N(t)^{\gamma}$, with $\gamma > 1$, over time.
- Graphs densify (!)
- Larger exponent, denser graph (!)

# Subgraph Densification laws



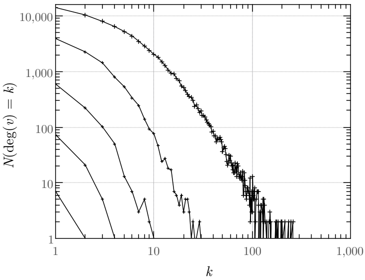

$$\bar{d} = \frac{e}{n} \qquad \rho = \frac{2e}{n(n-1)} \sim \frac{e}{n^2}$$

# DENSIFICATION LAWS (KLEINBERG, LESKOVEC, FAOLUTSOS)



# Results of Dorogovtsev and Mendes

- Accelerated growth of networks
- $\bar{k} \propto t^a$
- $P(k,t) \sim t^{-z} k^{-\gamma}$
- Scaling relations for accelerated growth
- The exponent changes in time in the Last.fm network:(



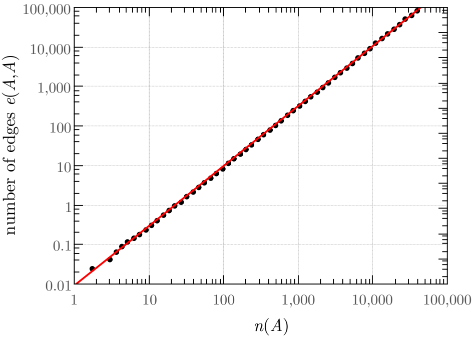
# Edges sampled from graphs - Pedarsani et al.

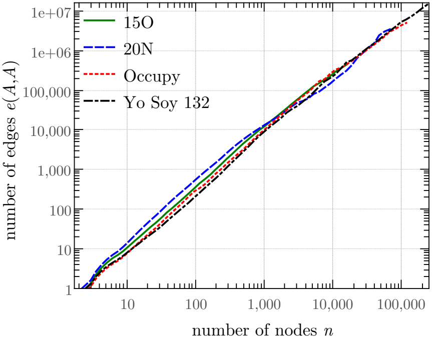
- Densification arising from sampling fixed graphs
- Fixed graph with power-law degree distribution
- Uniform sampling from the edges
- Direct relation between the exponent of the degree ditribution and the densification exponent.

# Real-world systems with background networks

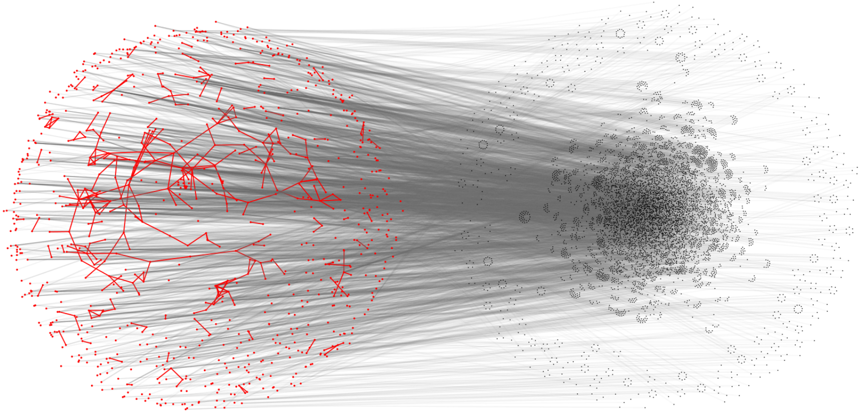
- Last.fm users scrobbling a given artist
- Retweeters of a message
- Physical world friendship and contacts appearing on a social network service (LinkedIn, Facebook, etc.)
- Physical world topics appearing on the Web
- Plenty of questions that the previous models can not explain

# SUBGRAPH DENSIFICATION LAWS

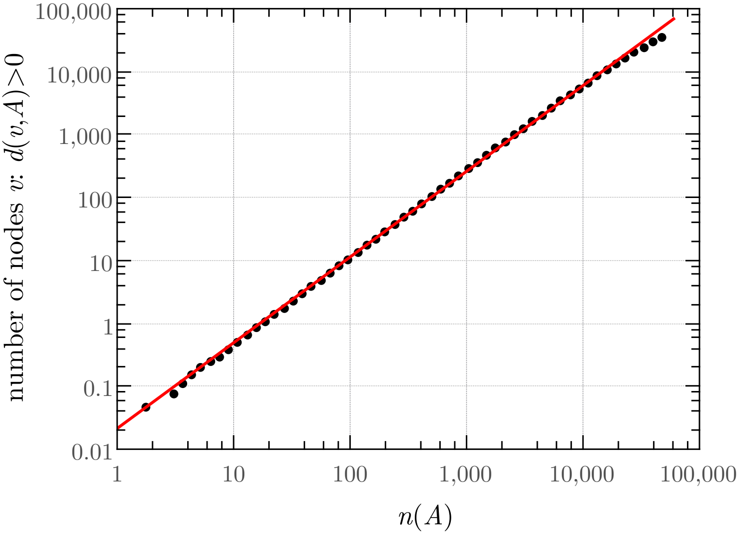


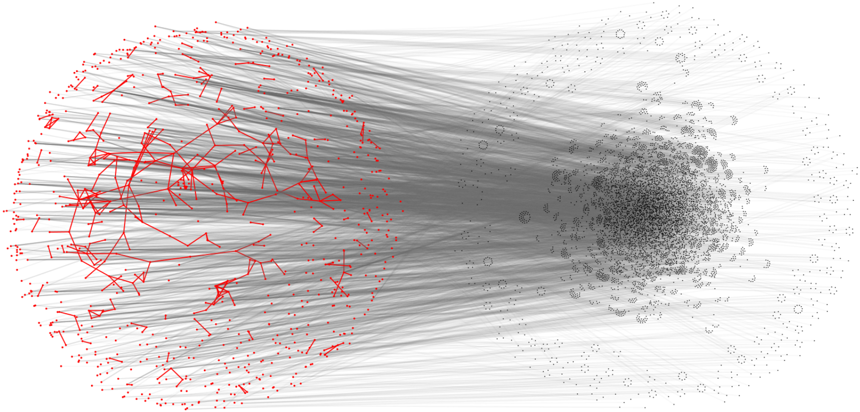
# Non-isolated nodes



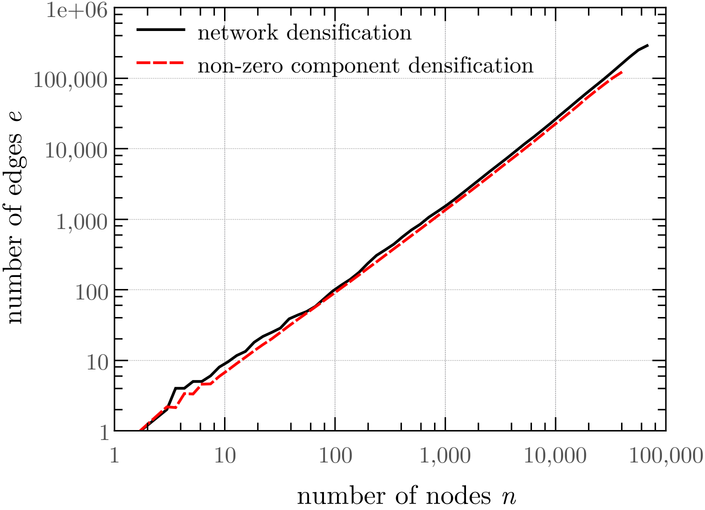
# Non-isolated nodes



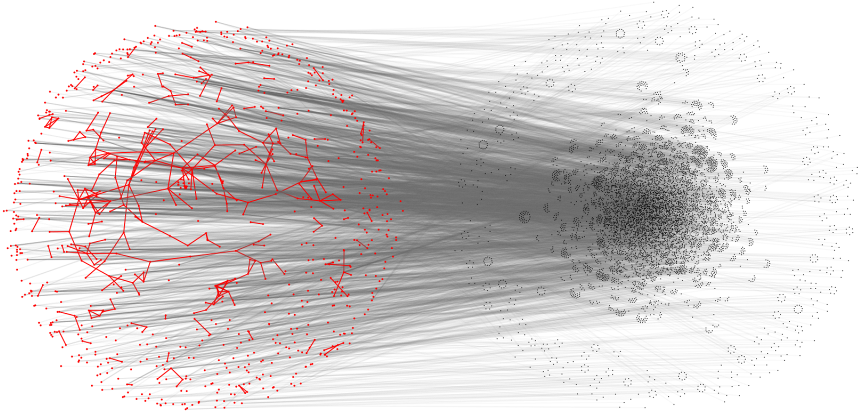
# NON-ISOLATED COMPONENT



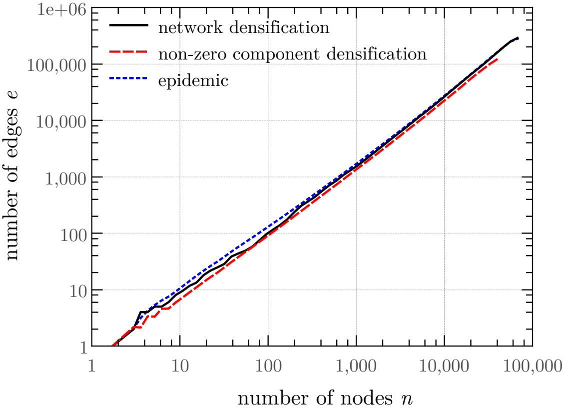
# Non-isolated component and Kleinberg densification



# Epidemic model

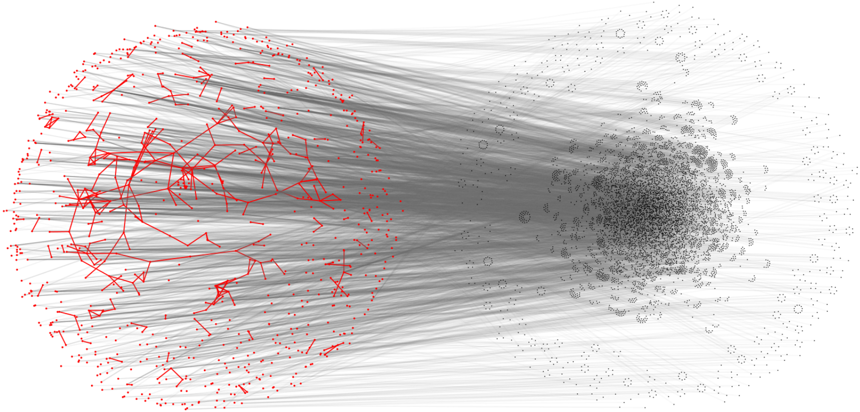


# Epidemic model

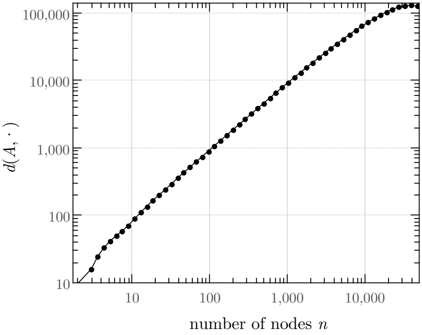


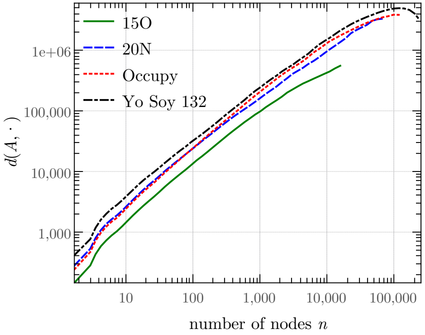

$$\beta \cdot \delta = \gamma$$

# Constant expansion



# Constant expansion



- “Densifying” community subgraphs with edge number following power law of node number
- Smaller subgraphs have higher *relative* density compared to a random subgraph of the same size
- This difference however vanishes with the community growth, the subgraph “sparsifies”
- Power law fraction of nodes with at least one edge within the community, with exponent greater than one
- Relation of exponents: $\beta \cdot \delta = \gamma$
- Information spreading over a network and the dynamic growth of the network are similar and closely related processes
- The network itself can be considered as a community in a hidden social network