

Network Structure & Privacy:

The Re-Identification Risk In Graph Data

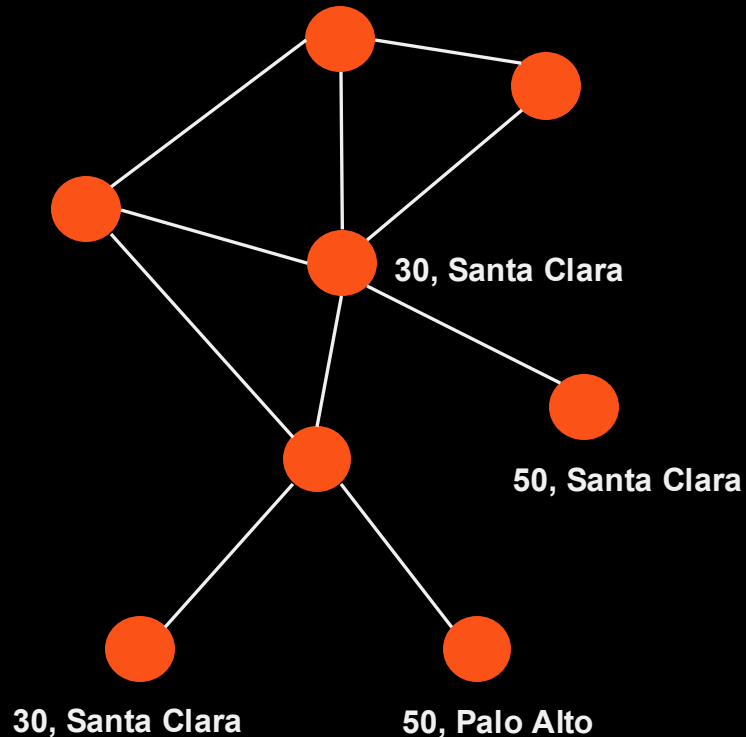
PEPR '25 – June 9th 2025

Daniele Romanini



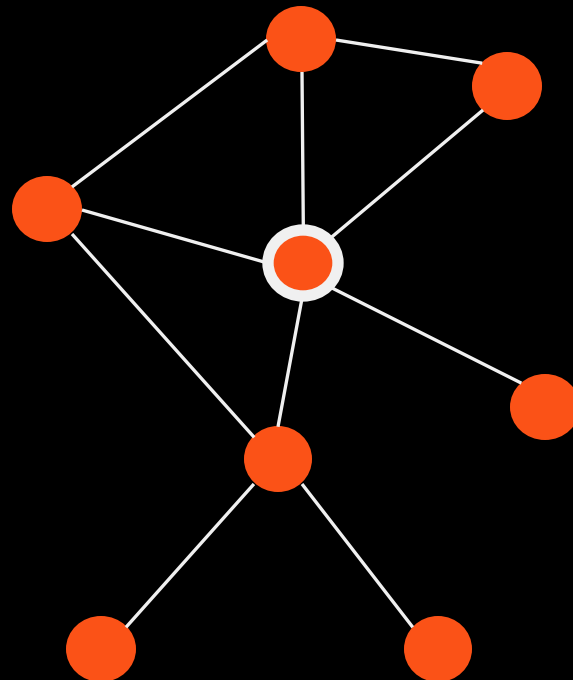
How unique are we?

| SSN | Age | City | ... |
|-------------|-----|-------------|-----|
| 123-45-6789 | 30 | Santa Clara | ... |
| 891-54-6788 | 50 | Palo Alto | ... |
| 987-65-4329 | 50 | Santa Clara | ... |
| 887-65-4320 | 30 | Santa Clara | ... |
| ... | ... | ... | ... |



How unique are we?

| SSN | Age | City | ... |
|-------------|-----|-------------|-----|
| 123-45-6789 | 30 | Santa Clara | ... |
| 891-54-6788 | 50 | Palo Alto | ... |
| 987-65-4329 | 50 | Santa Clara | ... |
| 887-65-4320 | 30 | Santa Clara | ... |
| ... | ... | ... | ... |



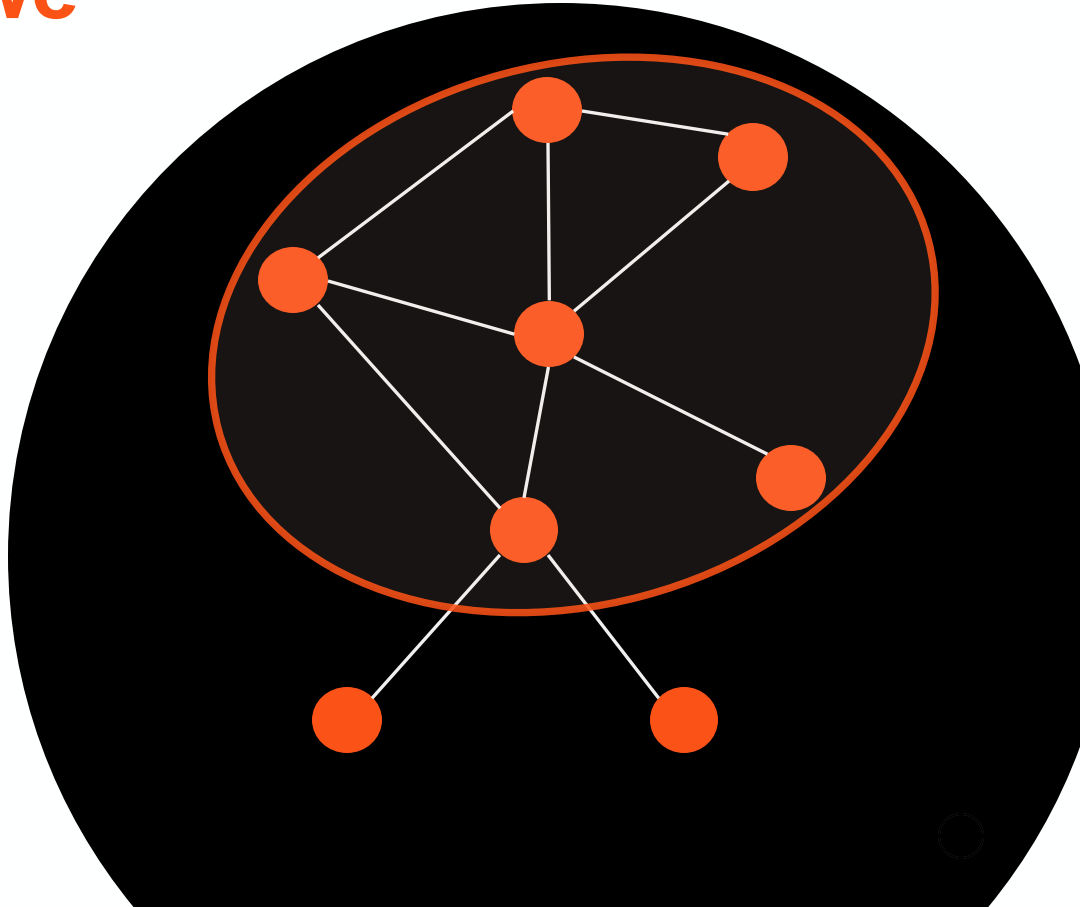
How unique are we based on our social circle?



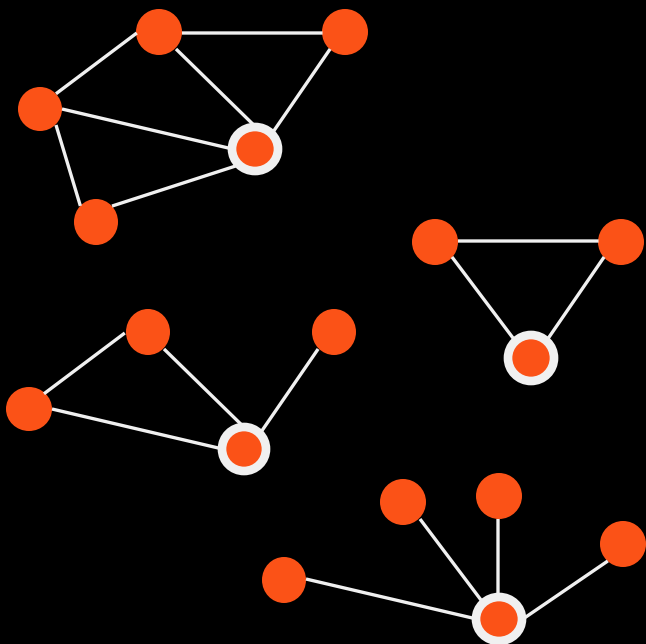
How many friends we have
(degree)



Connections between them
(neighborhood)



Studying the uniqueness of neighborhoods

**01**

Number of nodes
(network size)

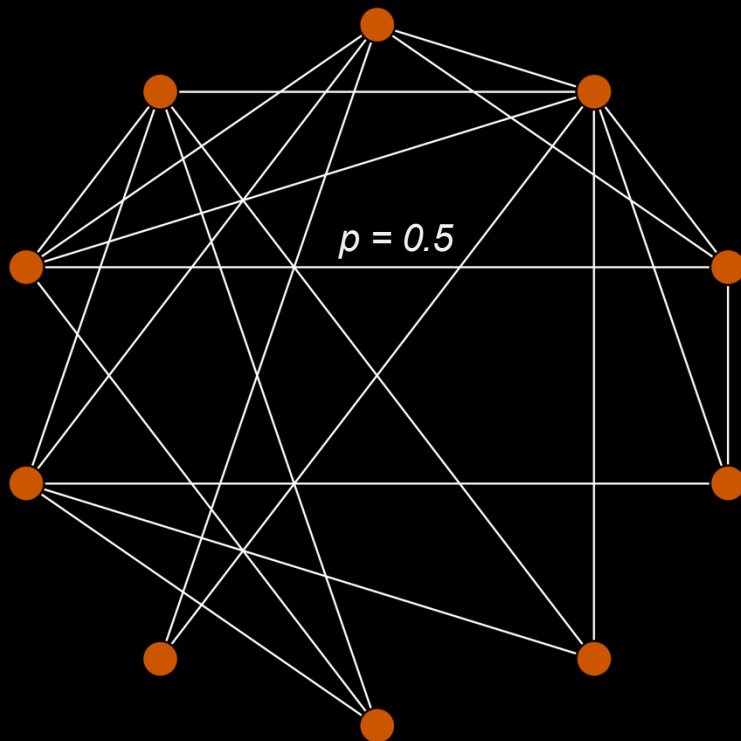
02

**Average amount of
connections per node**
(average degree)

03

Network structure
*(random, dense, sparse,
local communities, etc...)*

What if everything was random?



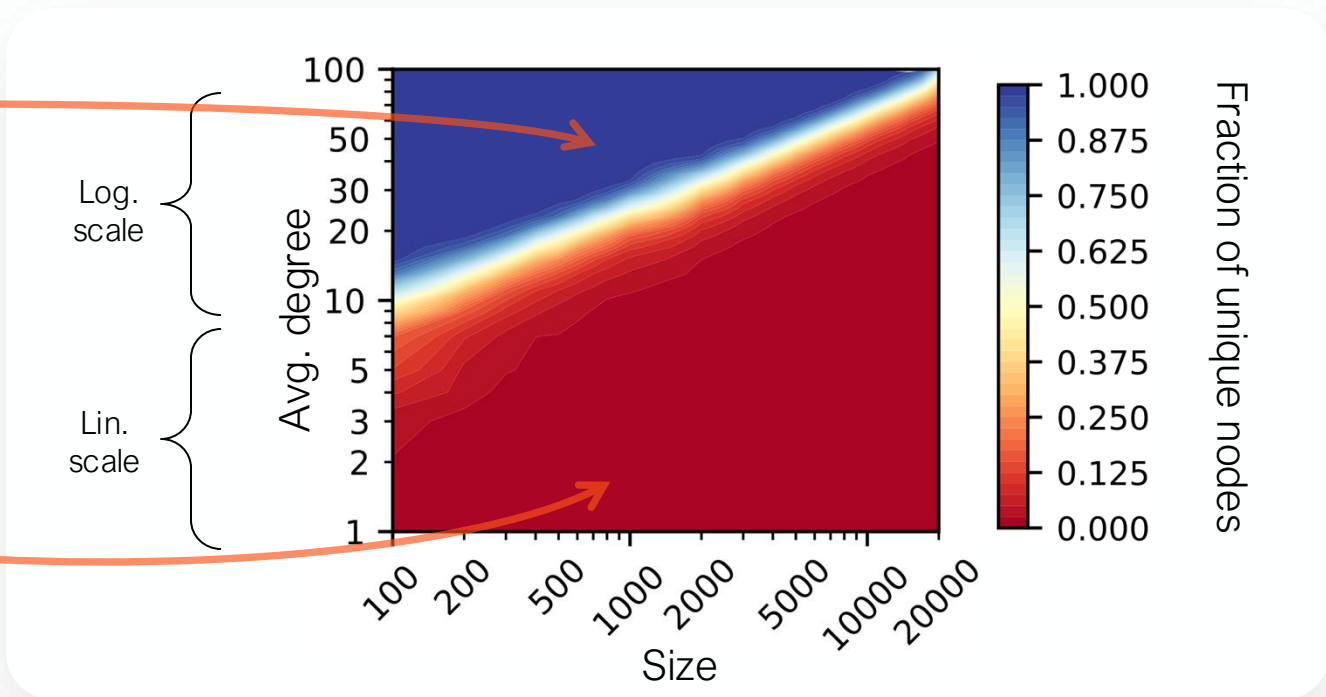
Erdős, P.; Rényi, A.,
"On random graphs "

Uniqueness map: size and average degree

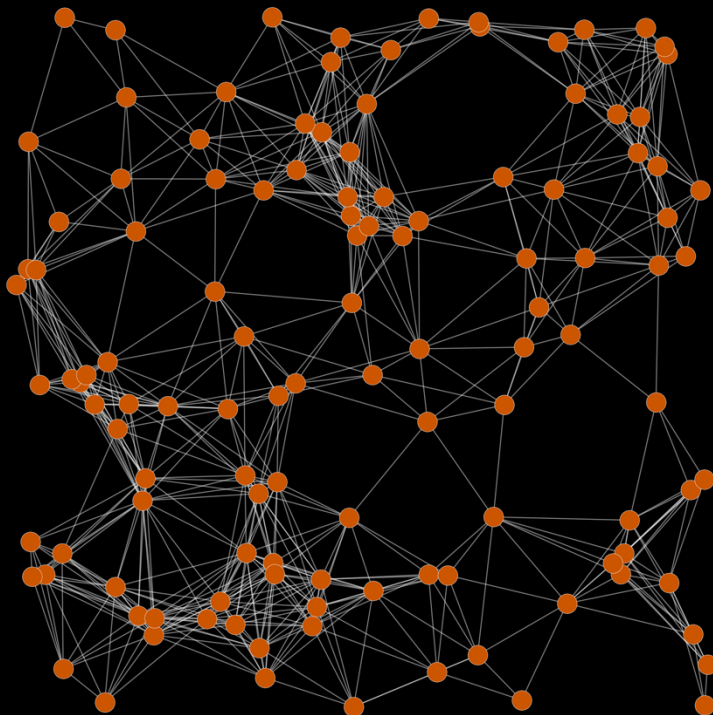
Romanini et al.,
Privacy and Uniqueness of
Neighborhoods in Social Networks

All
Unique
nodes
 $U = 1$

No
Unique
nodes
 $U = 0$



Adding more structure: Random Geometric Graph



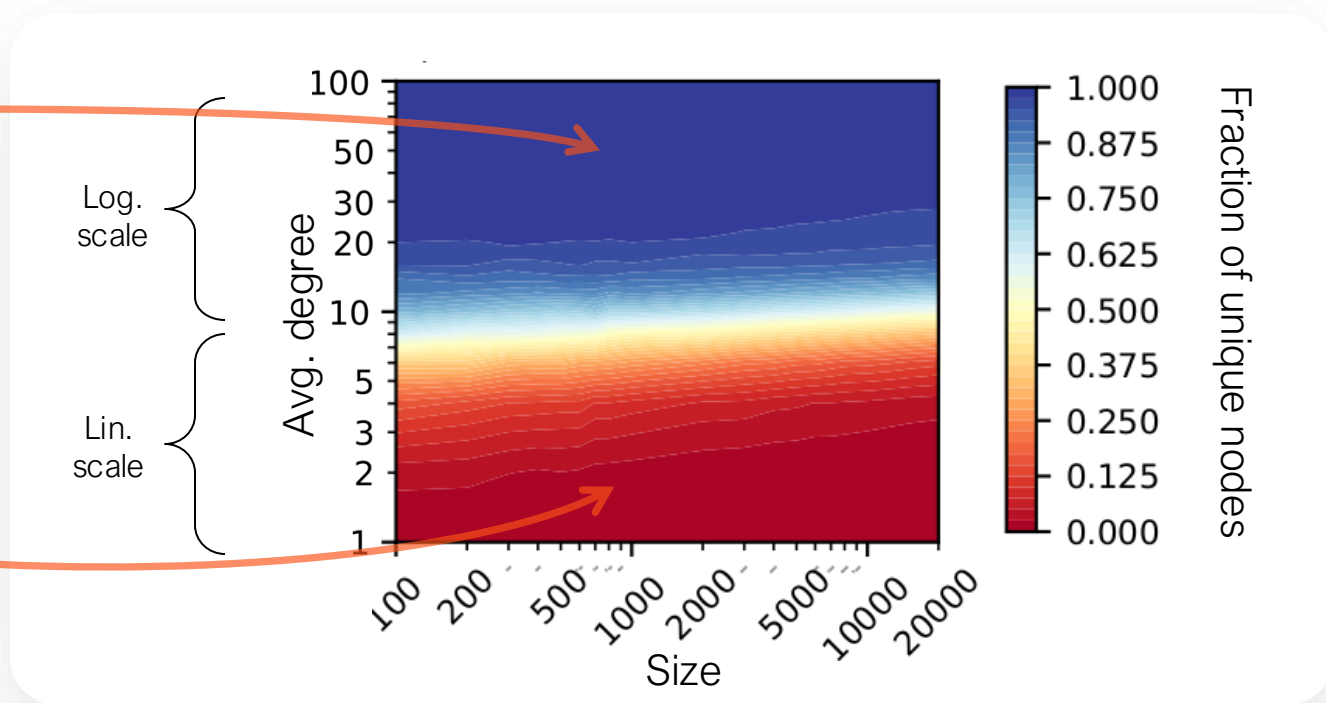
*M. Penrose,
"Random geometric graphs"*

Adding more structure: uniqueness map

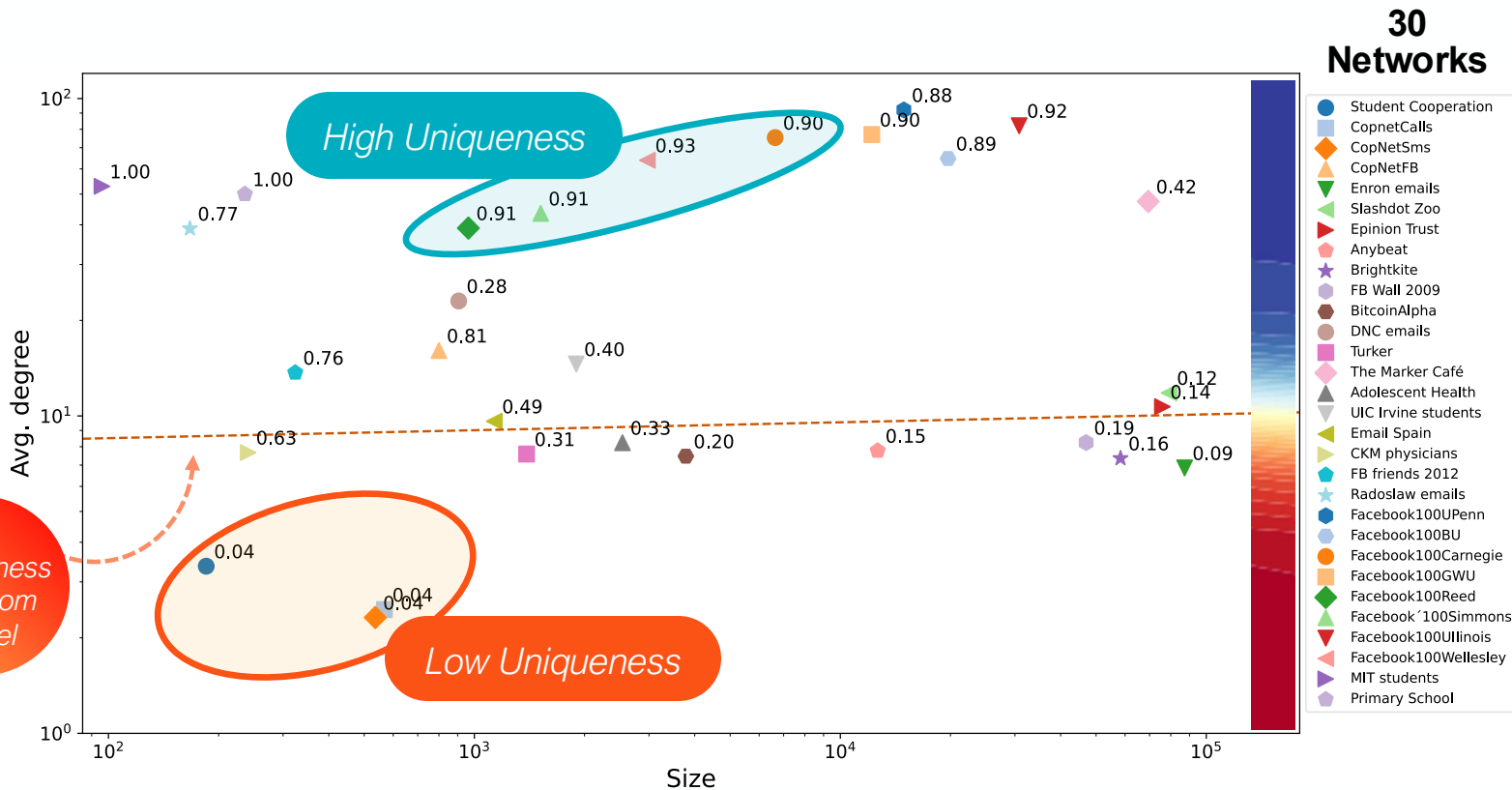
Romanini et al.,
Privacy and Uniqueness of
Neighborhoods in Social Networks

All
Unique
nodes
 $U = 1$

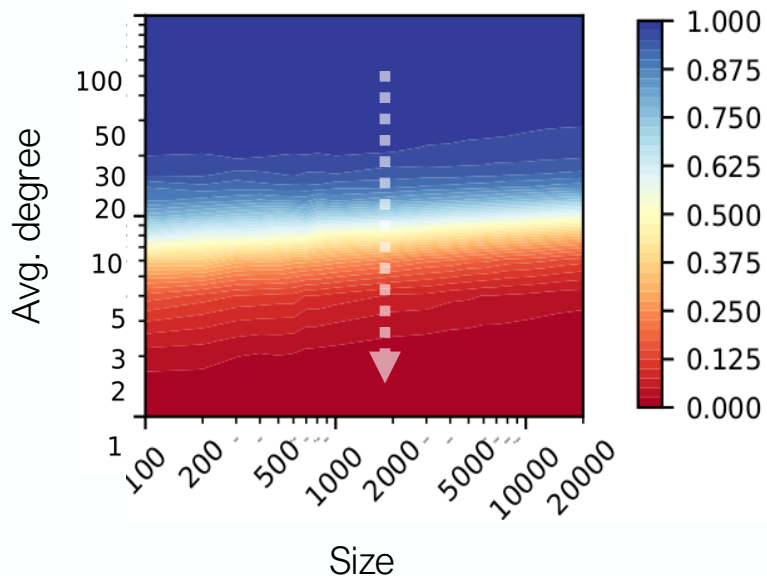
No
Unique
nodes
 $U = 0$



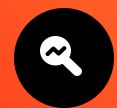
What about uniqueness in real data?



A simple approach: Edge Sampling?



Lowering avg. degree: lowering uniqueness



Error estimation and correction



Other attacks still possible e.g. existing links are there



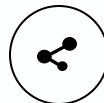
Other techniques: Differential Privacy (Node-DP, Edge DP)



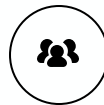
Take home



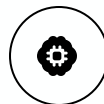
Uniqueness of neighborhoods
can be modelled and can be
a tool for risk assessment



Average degree > 10: danger!
The more the friends,
the higher the risk



Hiding yourself in the crowd
does not work;
local structure counts!



Anonymization techniques
should take structure
into account

Thank You

For your time

Co-Authors: Mikko Kivelä, Sune Lehmann



Bonus: aggregating networks with multiple layers

