Scientific Modeling Computer Laboratory

Project: Time Evolving Networks

Fourth Bi-weekly Presentation

by: Ádám Gergely Szabó

Supervisors:
Péter Pollner & Gergely Palla

2022, Spring Semester

Previously

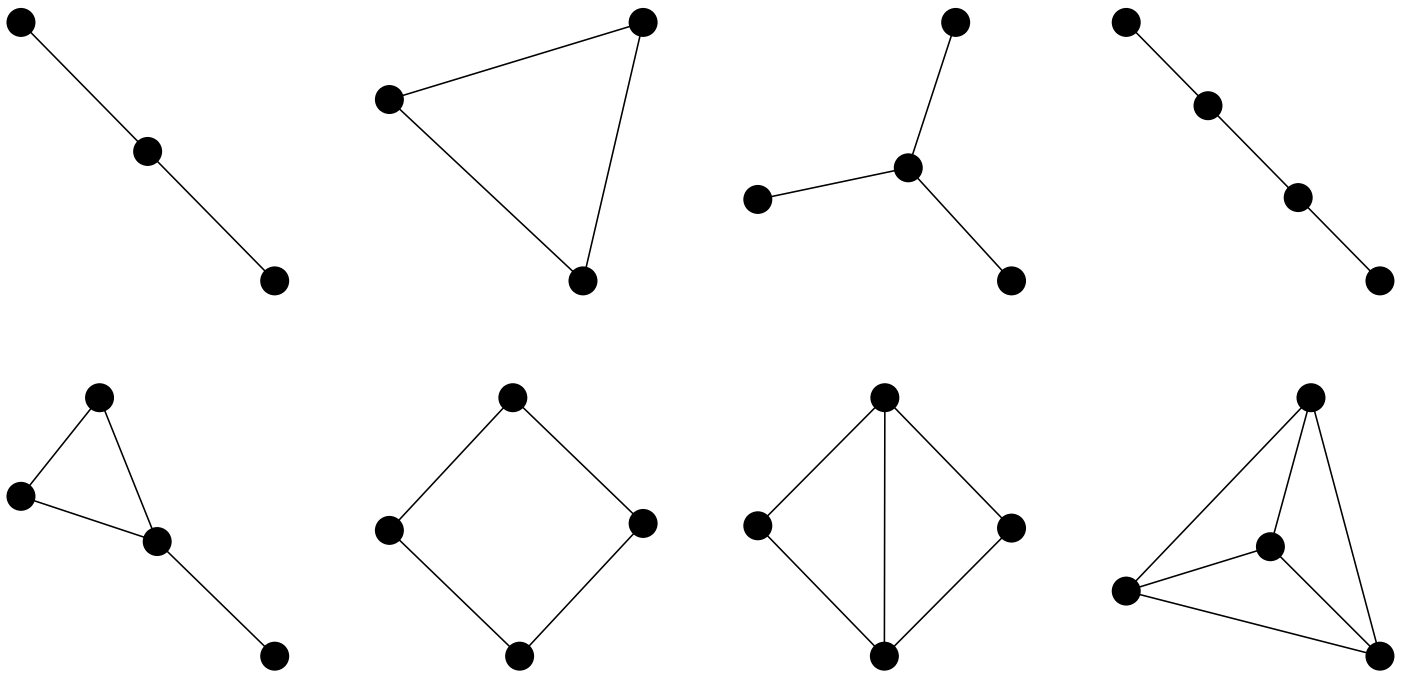
- What is MTMT?
  Hungarian Repository of Scientific Works
- How to acquire data?
  ReST API Queries
- What part of the data is needed?
  Authors for a given publication

Previously: Motifs I.

Motifs are little graphs with a given structure. The search for these graphlets are the enterance for group searching.

Unfortunately, the more nodes a motif has, the more expensive it becomes to find them.

Previously: Motifs II.

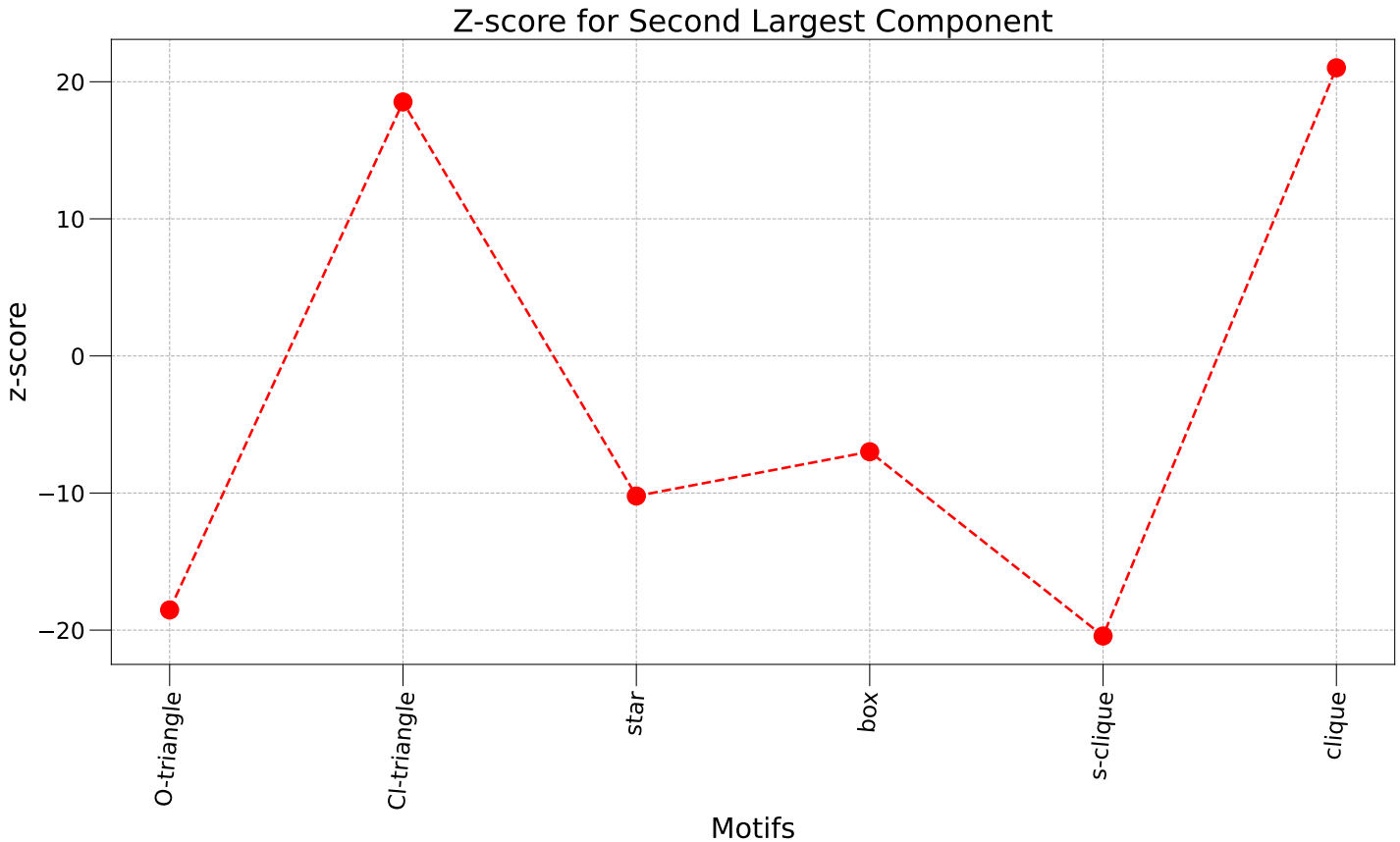


Z-score I.

Z-score is the way to measure the significance of a motif by comparing the original network to its randomised counterparts.

$$z = \frac{\langle m_i \rangle_g - \langle m_i \rangle_{rand}}{\sigma_{rand}} \tag{1}$$

Z-score II.


Z-score for Second Largest Component
z-score
20
10
0
−10
−20
O-triangle
Cl-triangle
star
box
s-clique
clique
Motifs

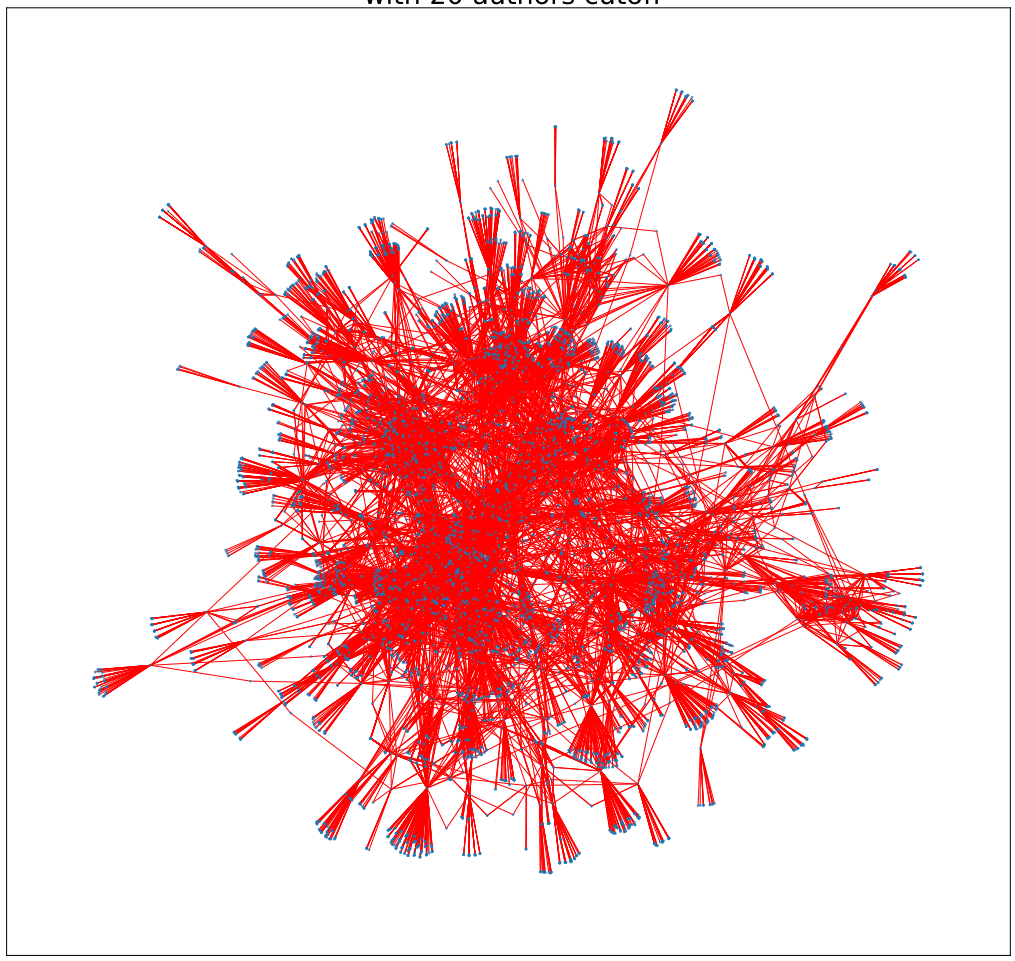
Problems with Motif Searching

The processing time of VF2 algorithm behind motif searching scales with the network size with $O(n^3)$
Solutions:

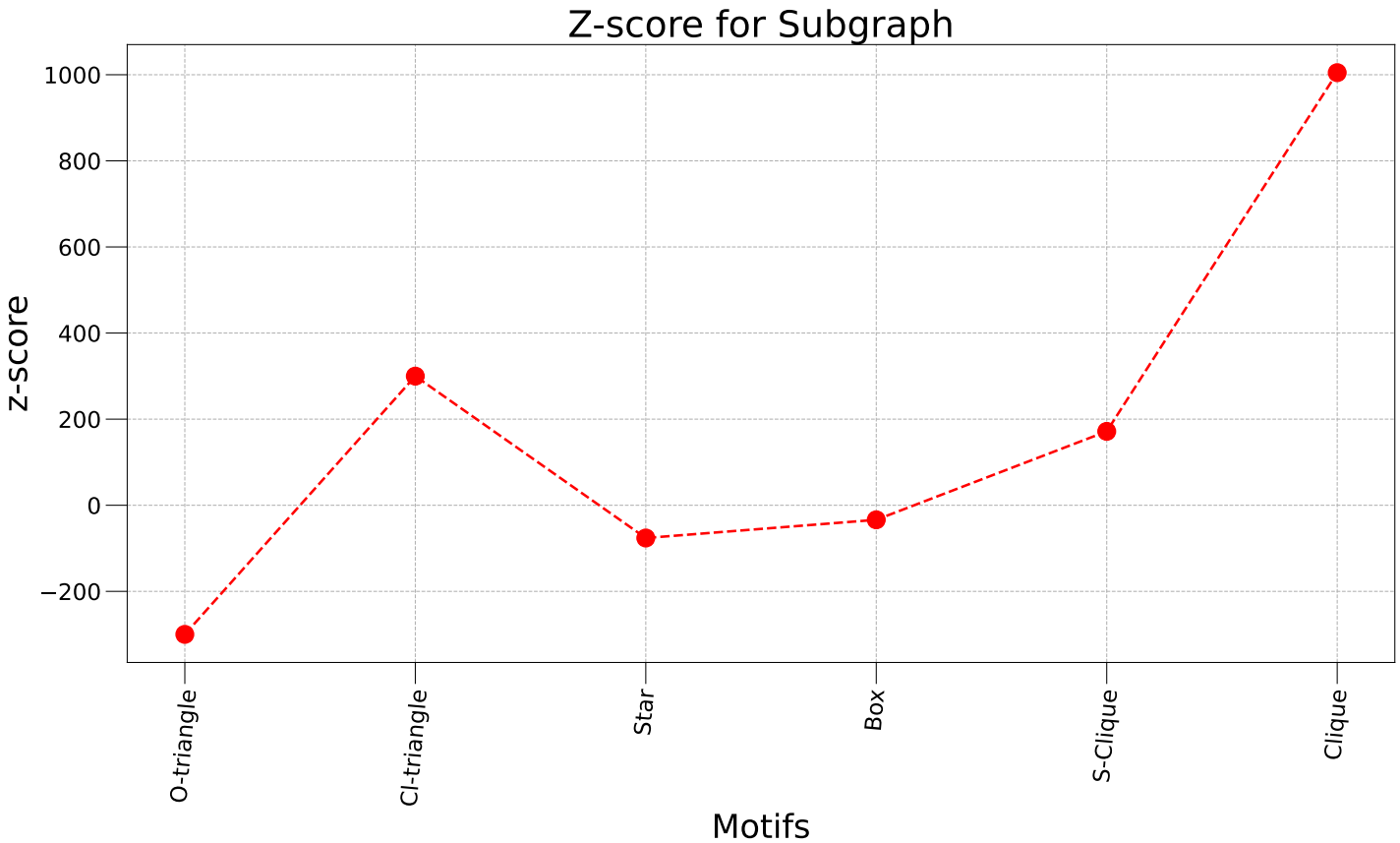
- Multiprocessing
- Subgraph with comparable size to the original graph

Subgraph I.

Graph Representation of the Subgraph of the Co-Author Network
with 20 authors cutoff



Subgraph II.


Z-score for Subgraph
1000
800
600
400
200
0
−200
Z-score
O-triangle
Cl-triangle
Star
Box
S-Clique
Clique
Motifs

Greedy Modularity I.

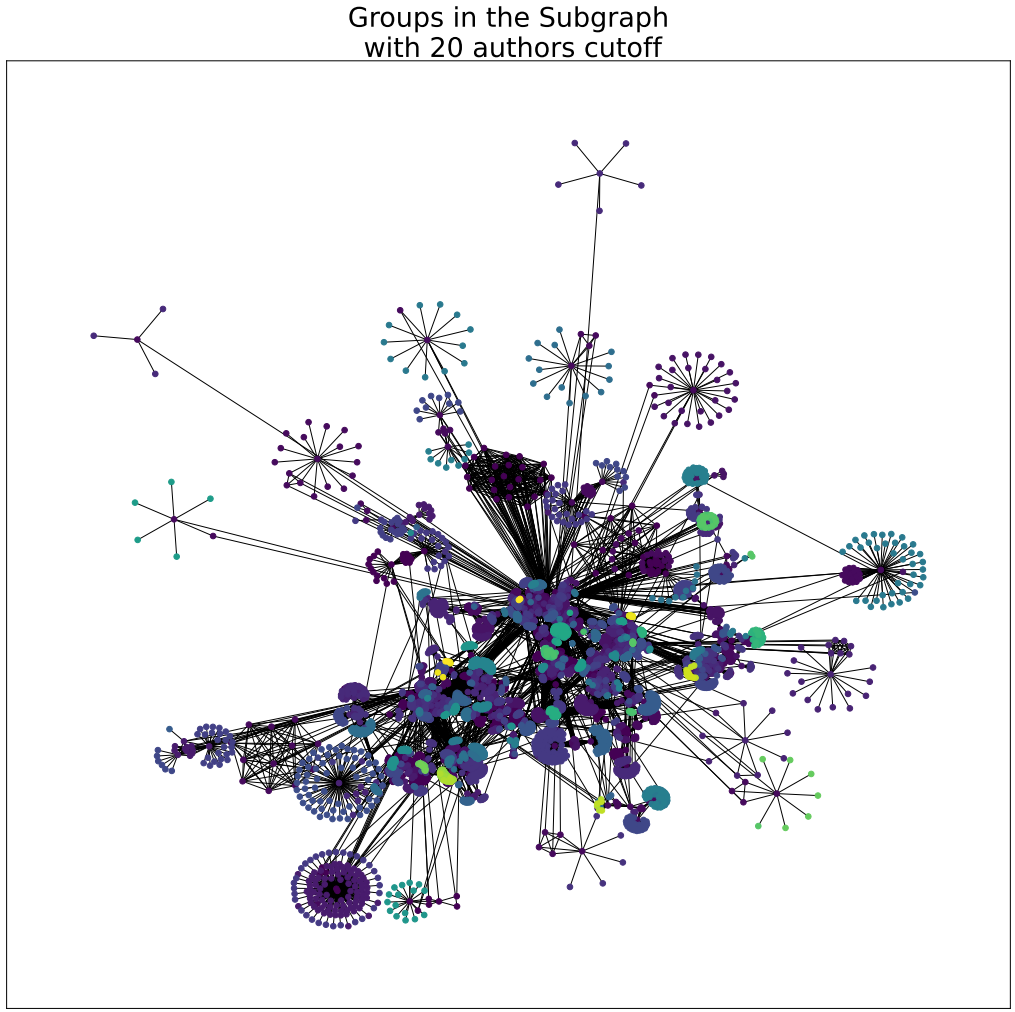
Greedy modularity is a hierarchical clustering based community searching method.

- Join nodes together depending on distance measure
- Measure modularity and do this until it reaches as maximum.

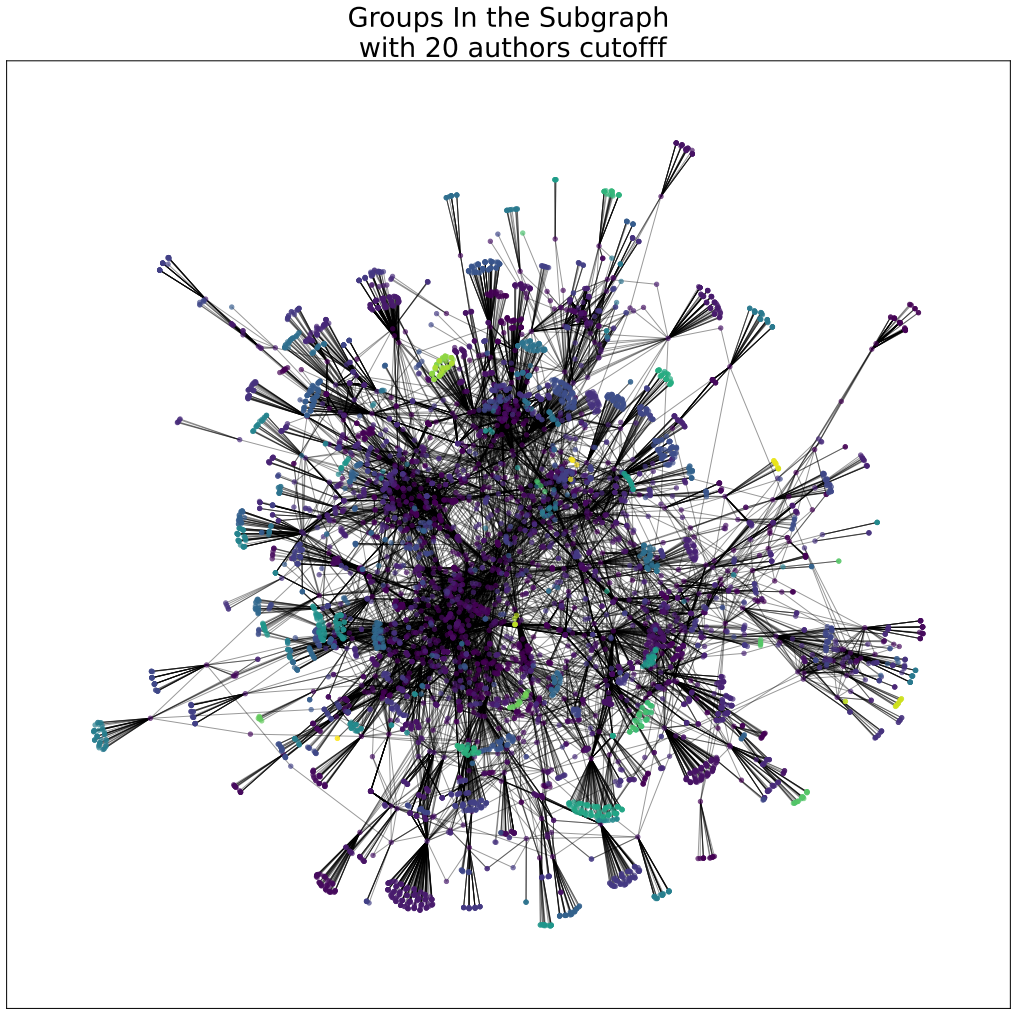
Pros	Cons
Much faster than motif finding	Every node will be part of a community

$$Q = \sum_{c=1}^{n} \left[\frac{L_c}{m} - \gamma \left(\frac{k_c}{2m} \right)^2 \right] \tag{2}$$

Communities I.



Communities II.



References

[1] Albert-László Barabási. "Network Science". In: *http://networksciencebook.com* (2012).

[2] Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. *Exploring network structure, dynamics, and function using NetworkX, in Proceedings of the 7th Python in Science Conference (SciPy 2008)*. 2008.

[3] Xiaoming Liu et al. *Co-Authorship Networks in the Digital Library Research Community*. 2005.

[4] Aaron Clauset, M. E. J. Neumann, and Cristopher Moore. "Finding community structure in very large networks". In: (2004).

Thank you for your attention!