Scientific Modeling Computer Laboratory

Project: Time Evolving Networks

Final Presentation

by: Ádám Gergely Szabó

Supervisors:
Péter Pollner & Gergely Palla

2022, Spring Semester

Description I.

- 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

Description II.

How to see time-evolution?

- Generate the network from beginning up to a given year
- Use a time interval to generate the network

One tracks connections ever made, other tracks activity.

Z-score and motifs I.

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

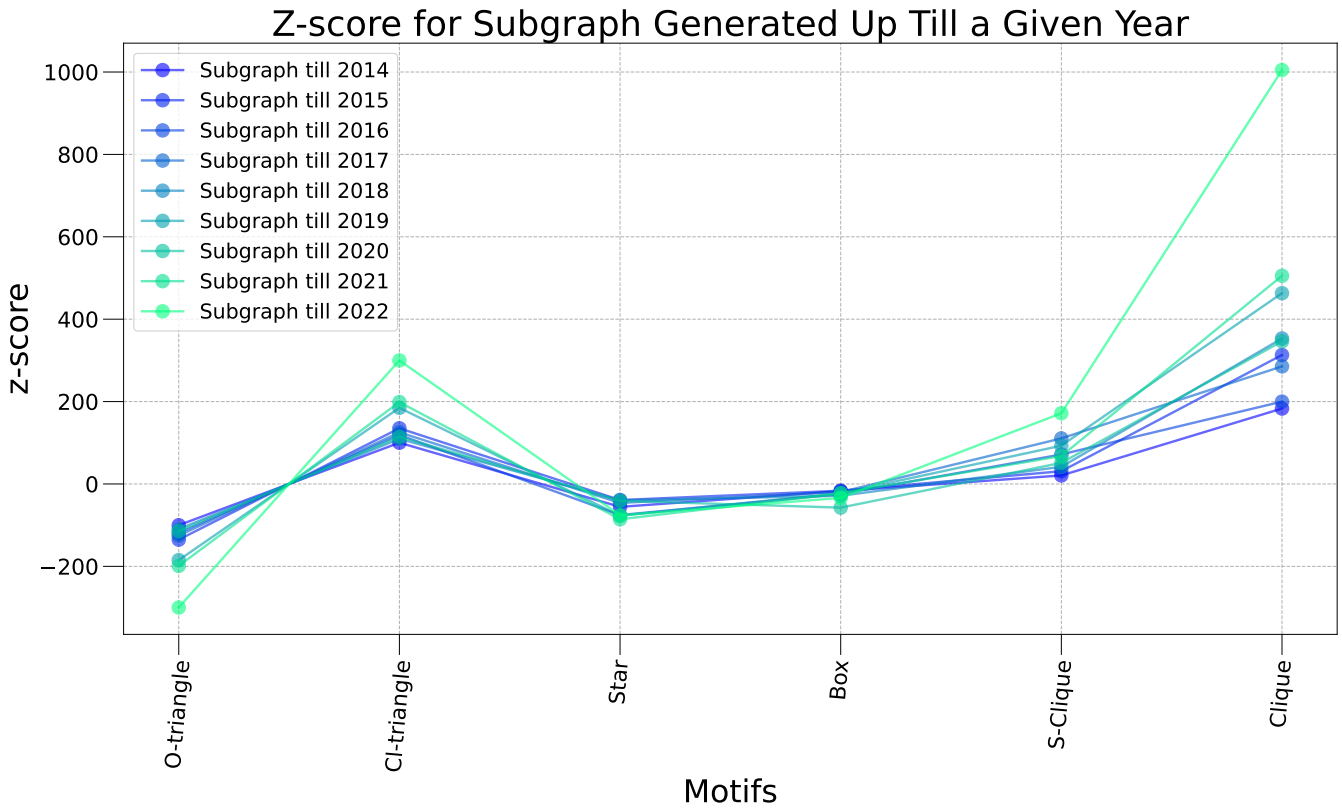
Unfortunately, the more nodes a motif has, the more expensive it becomes to find them.

Z-score and motifs II.

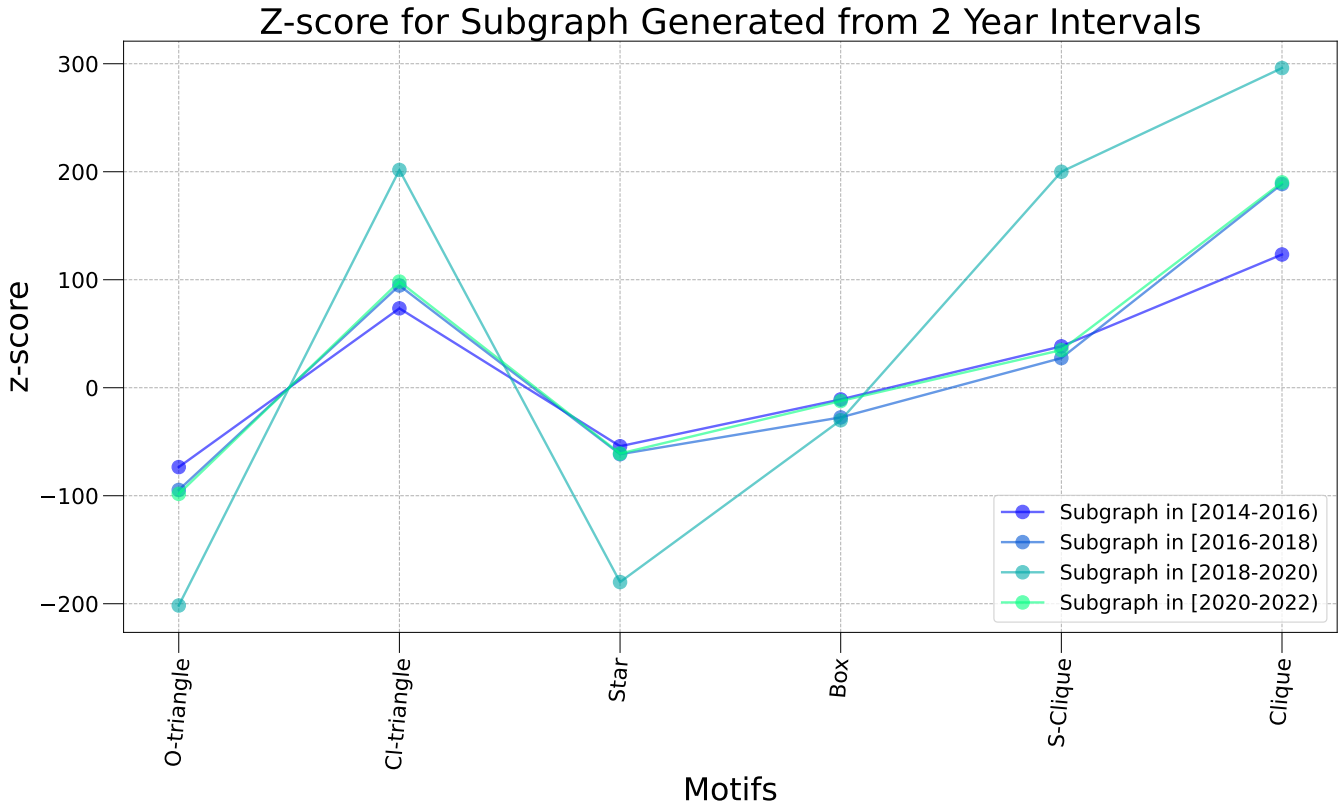
Z-score is the way to measure the significance of a motif by comparing the original network to its randomised counterparts.

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

Z-score and motifs III.



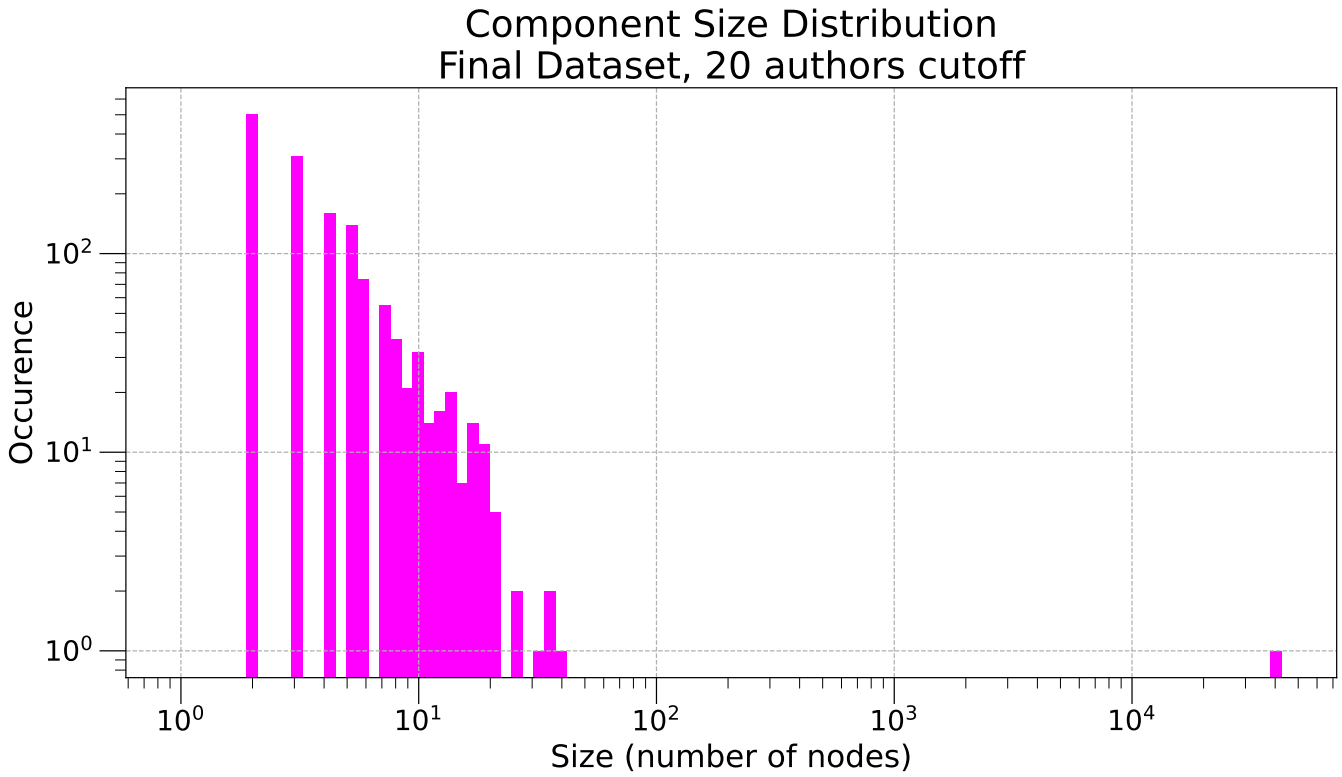
Z-score and motifs IV.



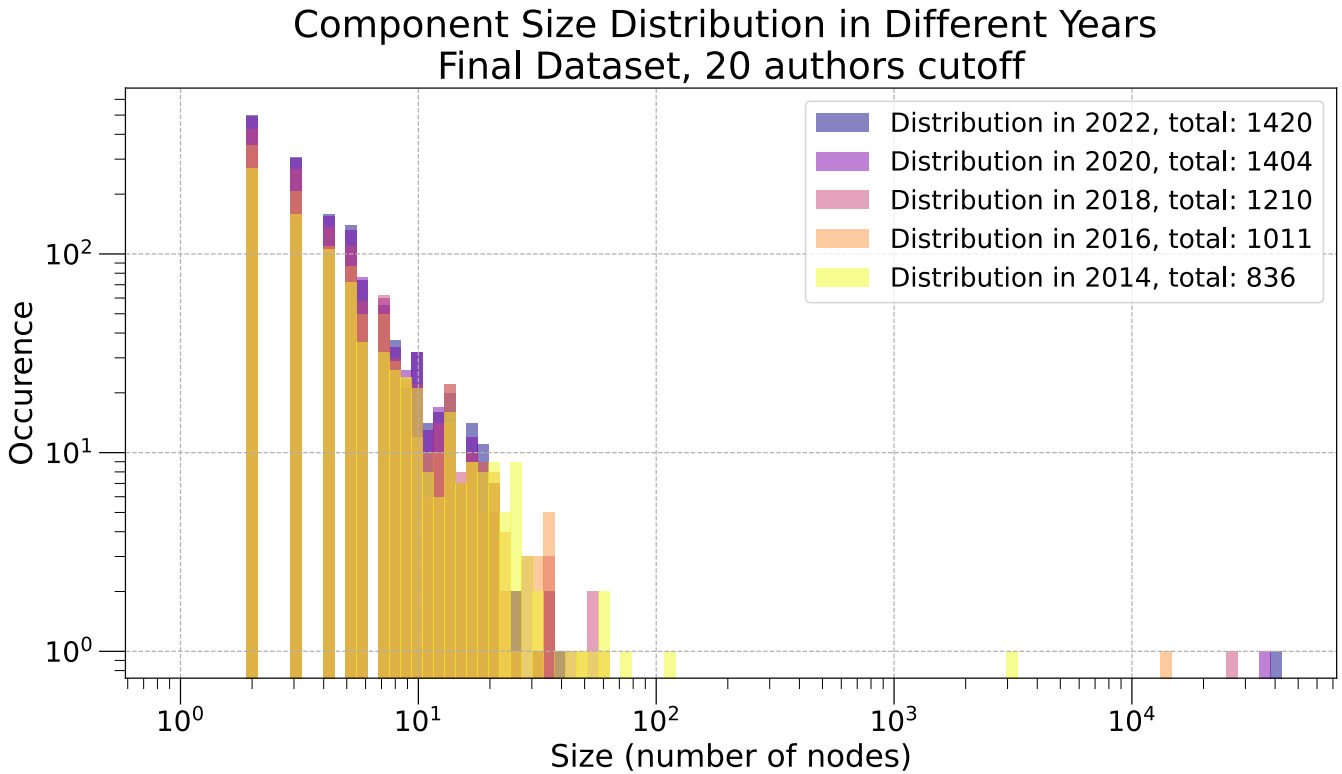
Components I.

Components are sets of nodes that are not linked to other sets of nodes. The Giant Component is a component which is comparable in size to the network.

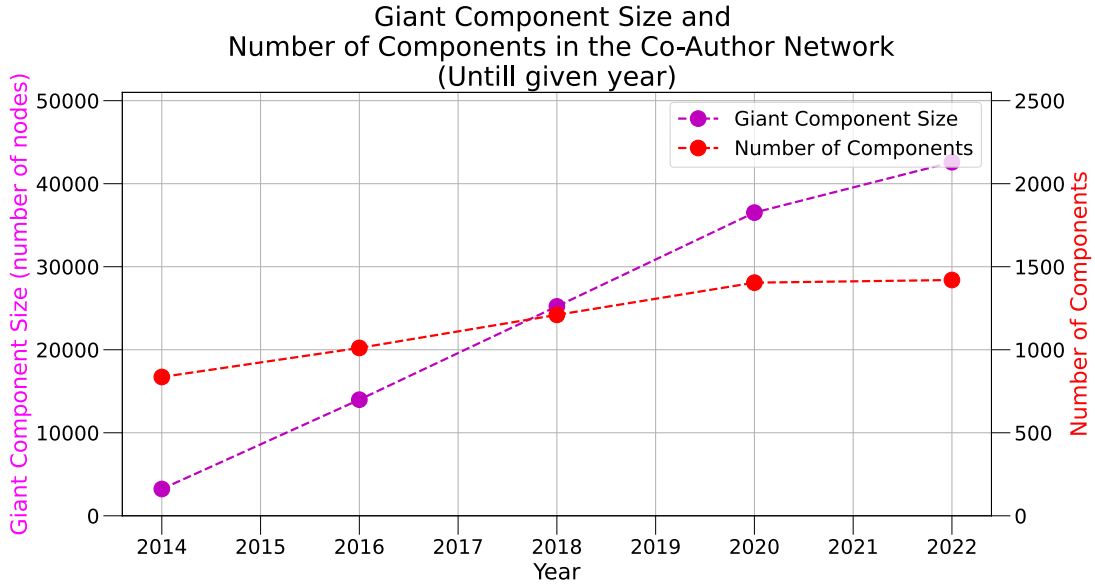
Components II.

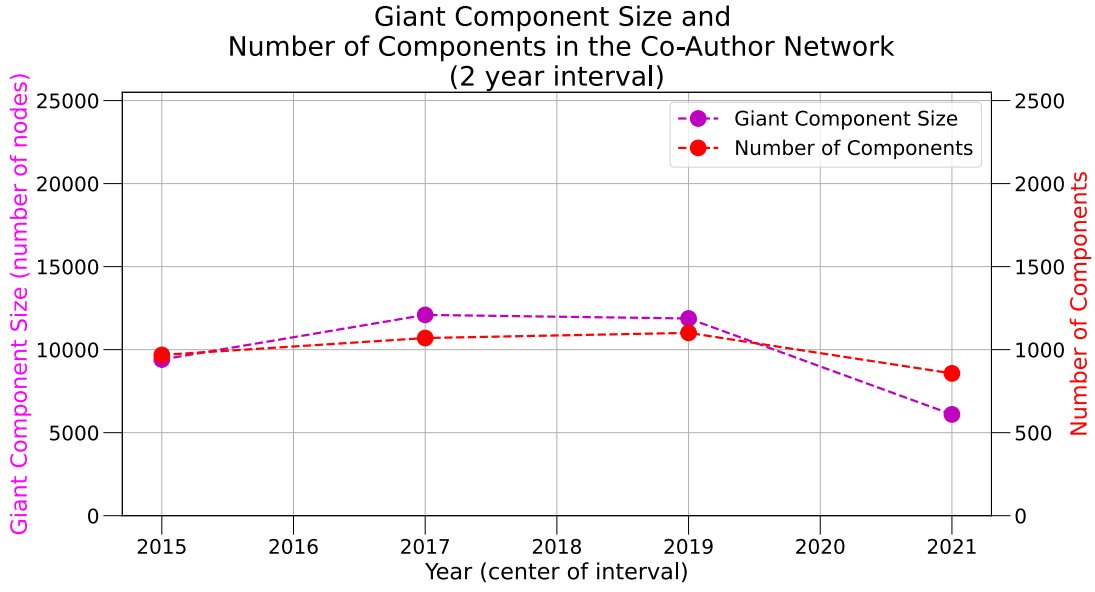


Components III.

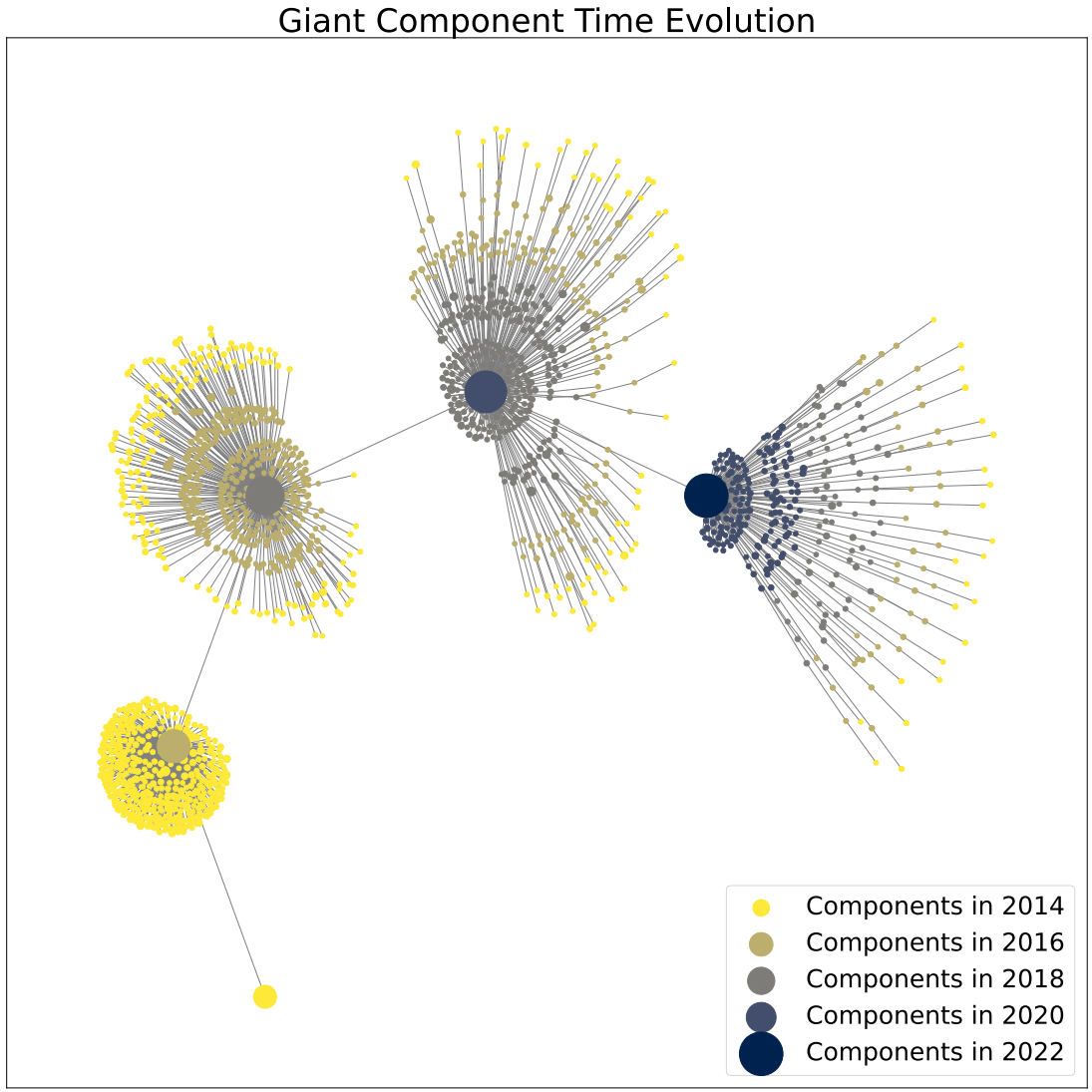


Components IV.

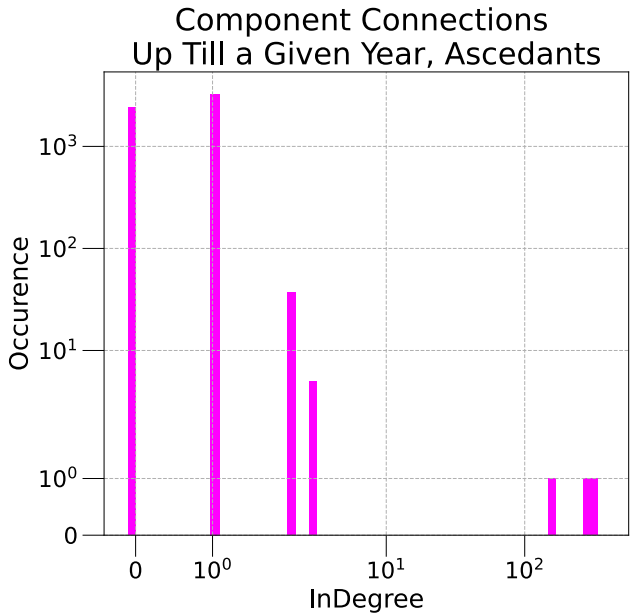


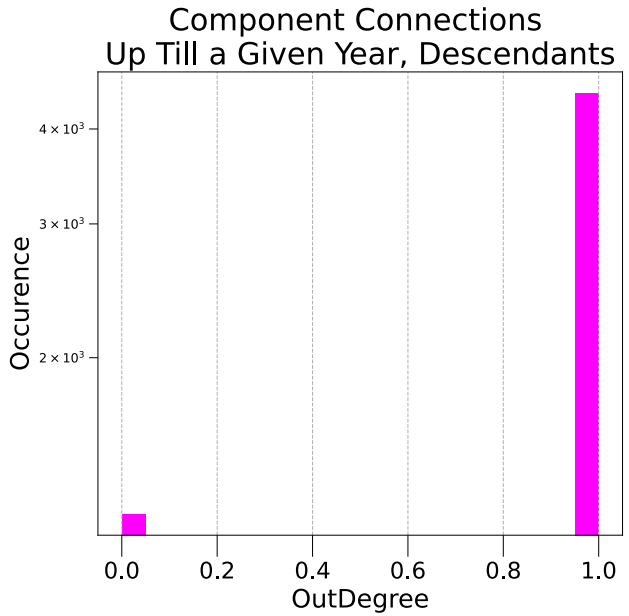


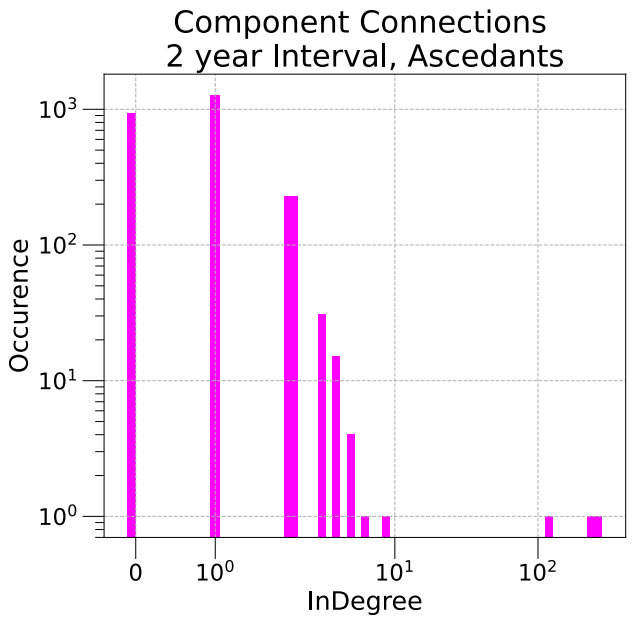
Components V.

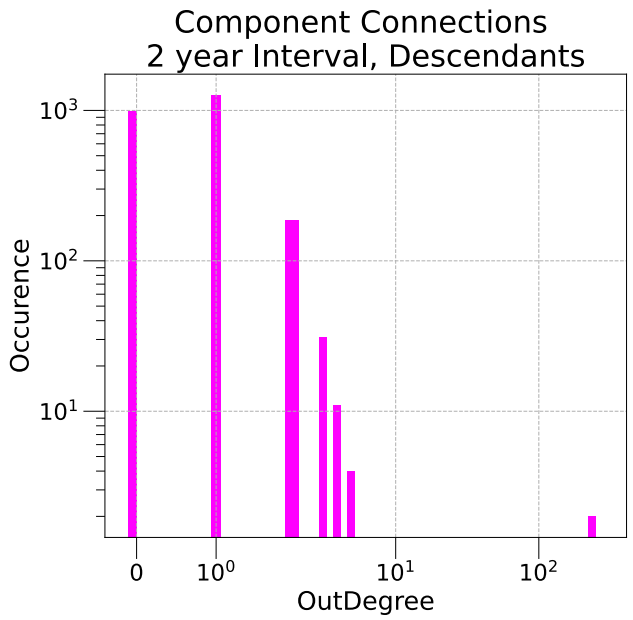


Components VI.









Greedy Modularity I.

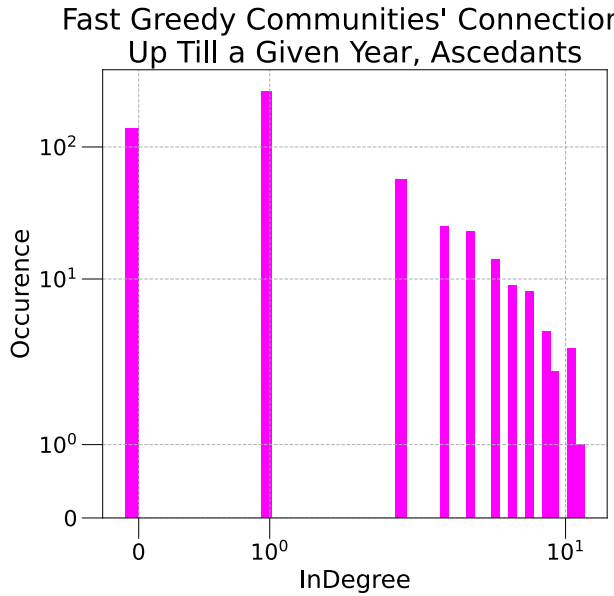
Greedy modularity is a hierarchical clustering based community searching method.

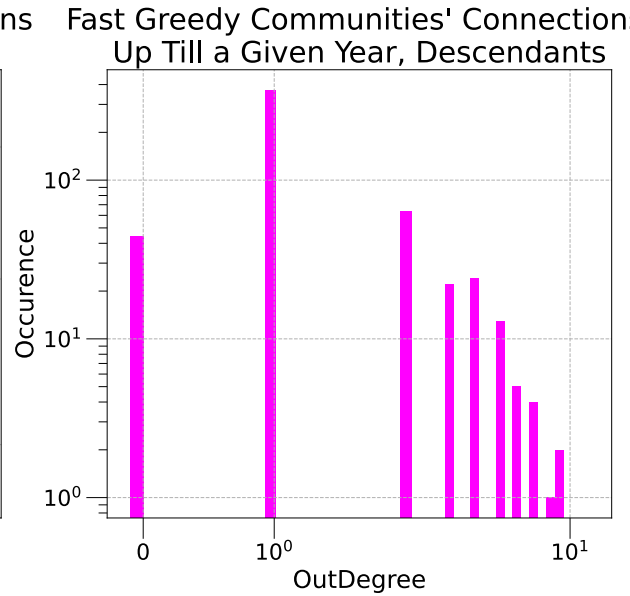
- Join nodes together with respect to the greatest increase in modularity
- Measure modularity and do this until it reaches as maximum.

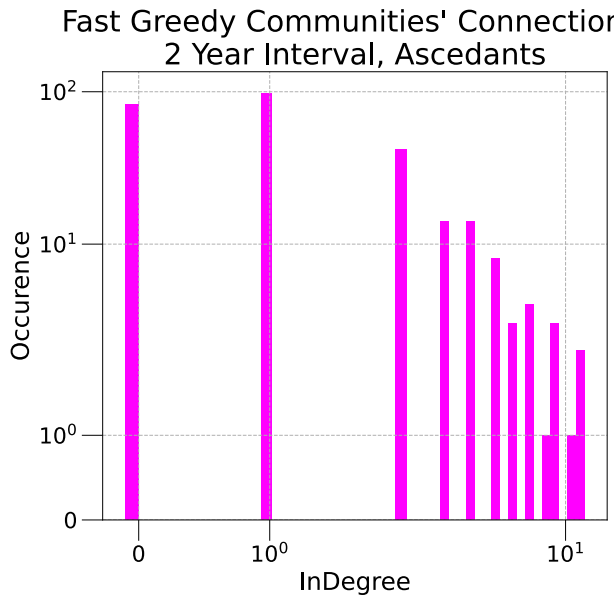
Pros	Cons
Much faster than motif finding	Every node will be part of a community

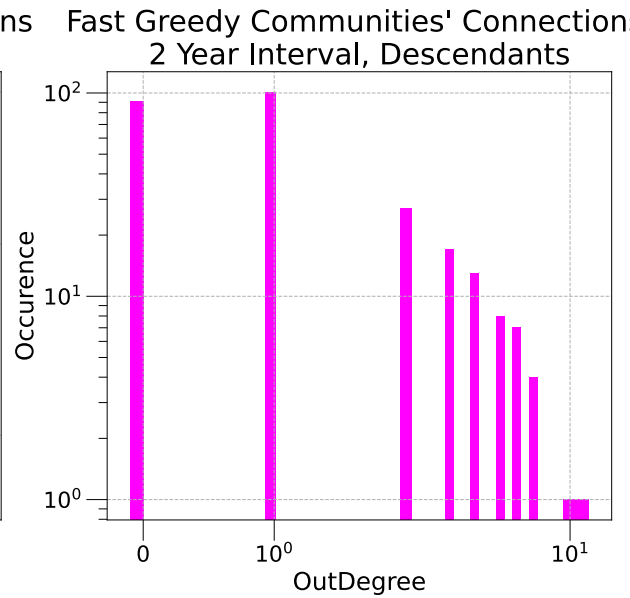
$$Q = \sum_{c=1}^{n} \left[\frac{L_c}{m} - \gamma \left(\frac{k_c}{2m} \right)^2 \right] \tag{2}$$

Greedy Modularity II.







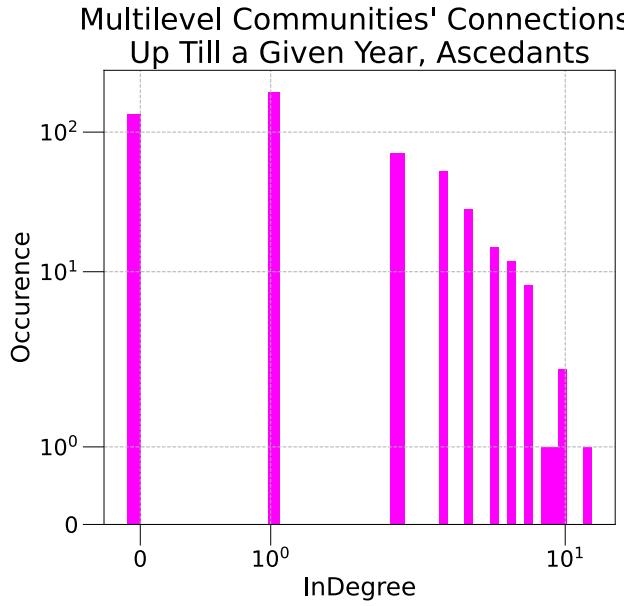


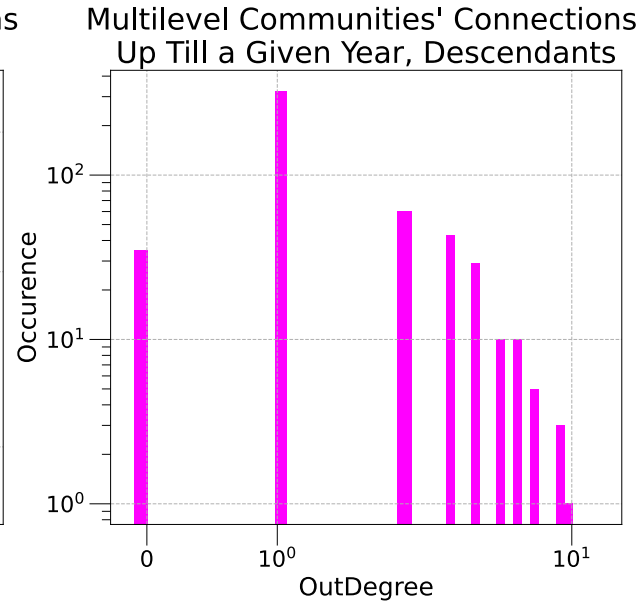
Multilevel (Louvain) I.

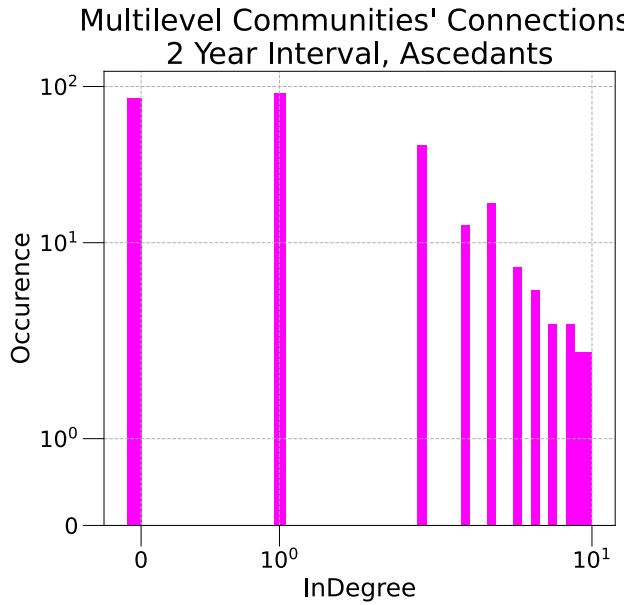
Multilevel is based on modularity optimization and community aggregation. It starts as the same as greedy modularity.

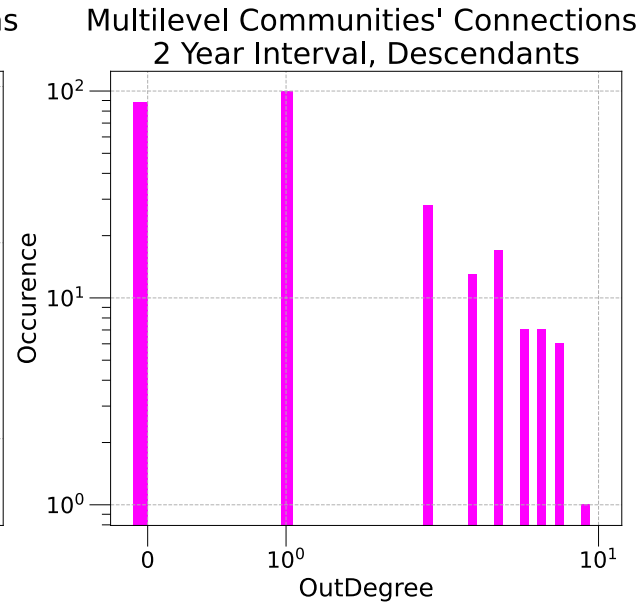
- Try to move nodes to another community, calculate the change in modularity and join the ones with the highest
- All nodes in the same community are merged into a giant node
- Repeat

Multilevel (Louvain) II.







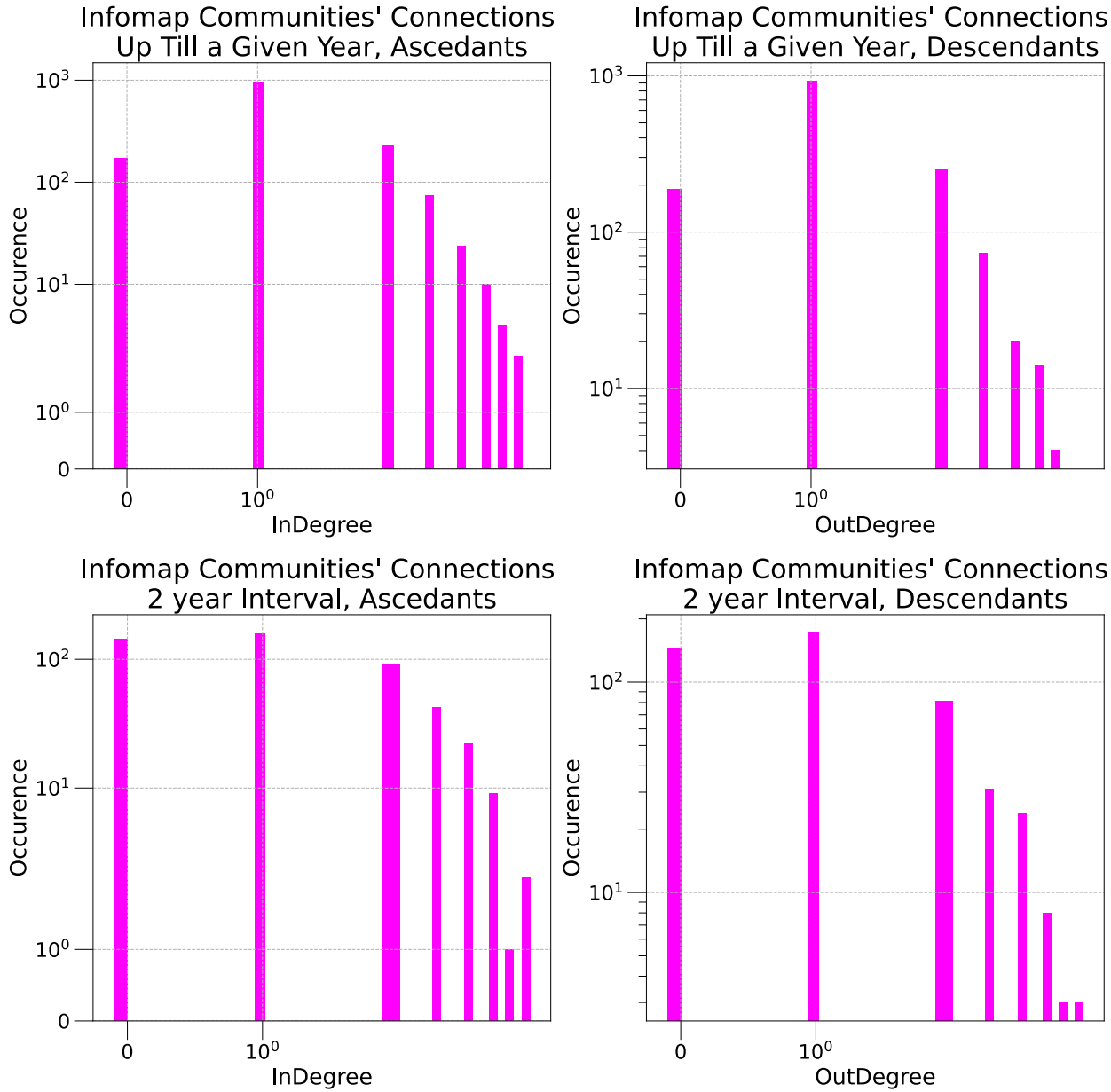


Infomap I.

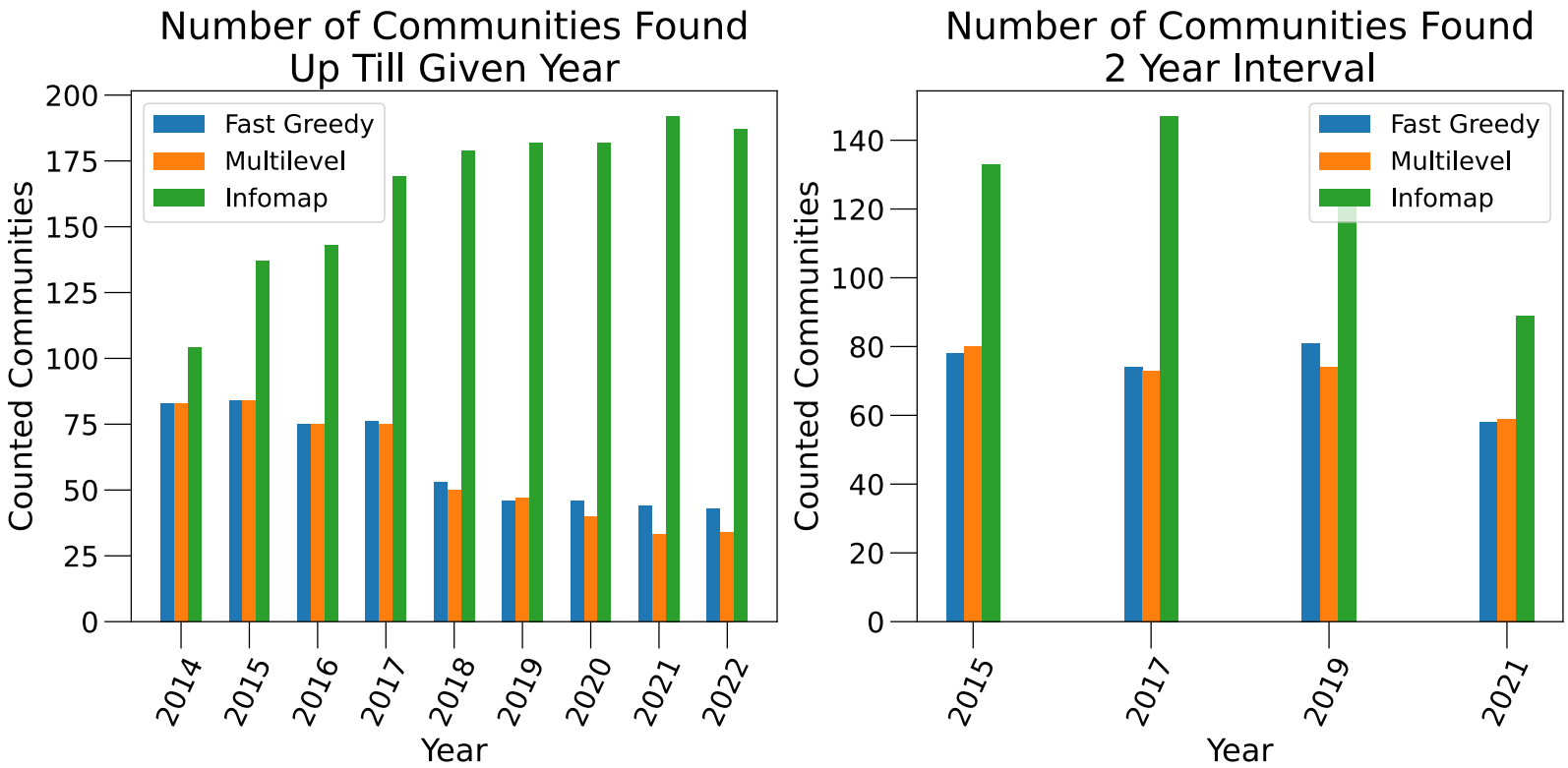
Infomap is vastly different from others before. It is based on the map equation, Shannon's source coding theorem and Huffman coding. The map equation takes the form as follows:

$$L(M) = qH(Q) + \sum_{m=1}^{n_m} p_o^m H(P_m) \tag{3}$$

Infomap II.


Infomap Communities' Connections
Up Till a Given Year, Ascedants
Occurence
InDegree
Infomap Communities' Connections
Up Till a Given Year, Descendants
Occurence
OutDegree
Infomap Communities' Connections
2 year Interval, Ascedants
Occurence
InDegree
Infomap Communities' Connections
2 year Interval, Descendants
Occurence
OutDegree

Differences I.



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).

[5] M. Rosvall, D. Axelsson, and C. T. Bergstrom. "The Map Equation". In: (2009).

Thank you for your attention!