Clique Community Persistence:

A Topological Visual Analysis Approach for Complex Networks

B. Rieck, U. Fugacci, J. Lukasczyk, H. Leitte

Technische Universität Kaiserslautern

Complex Networks

A ***network*** is a **complex system** consisting of **individuals** or **entities** connected by specific **ties** such as

- ***Personal Relationship***
- ***Shared Knowledge***

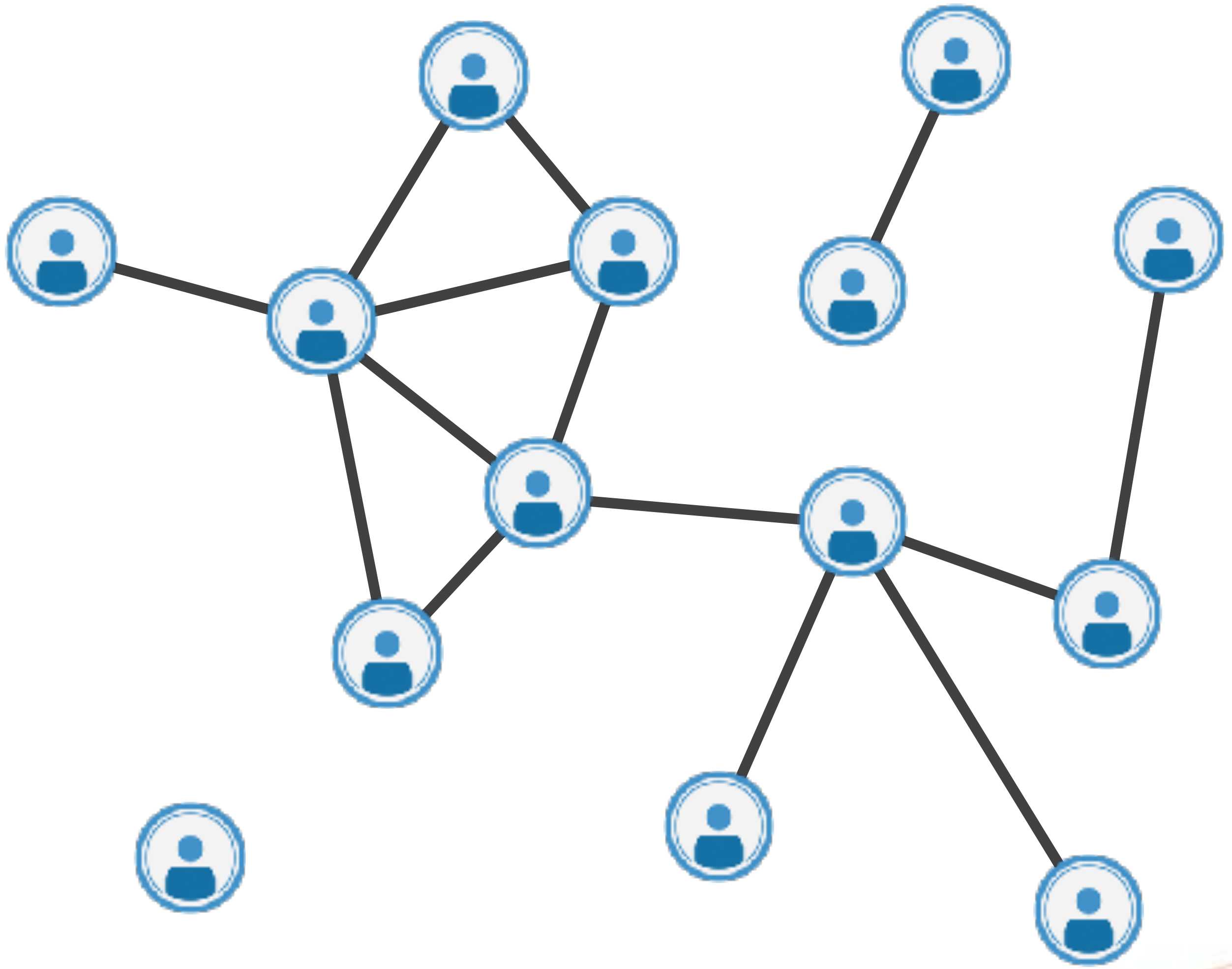
Several Examples:

- ***Social*** *Networks*
- ***Sensor*** *Networks*
- ***Biological*** *Networks*
- ***Collaborative*** *Networks*
- …

Complex Networks

A network can be represented by a **graph** ***G=(V, E)*** such that:

individuals	⇔	*nodes*
ties		*edges*
tie strength		*edge weight*



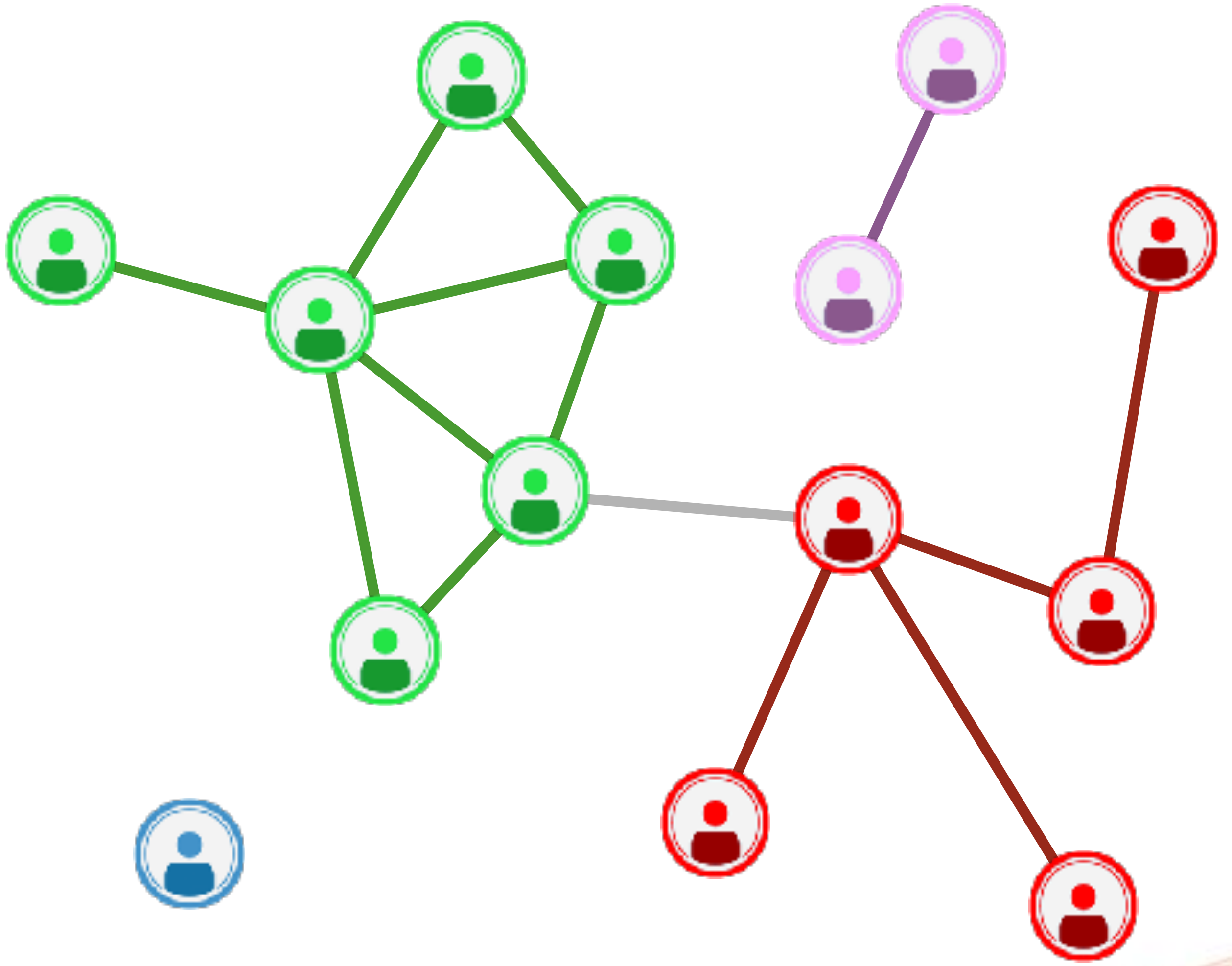
Complex Networks

A network can be represented by a **graph** ***G=(V, E)*** such that:

individuals	⇔	*nodes*
ties		*edges*
tie strength		*edge weight*

Network Analysis:

- *Retrieval of **global** structure*

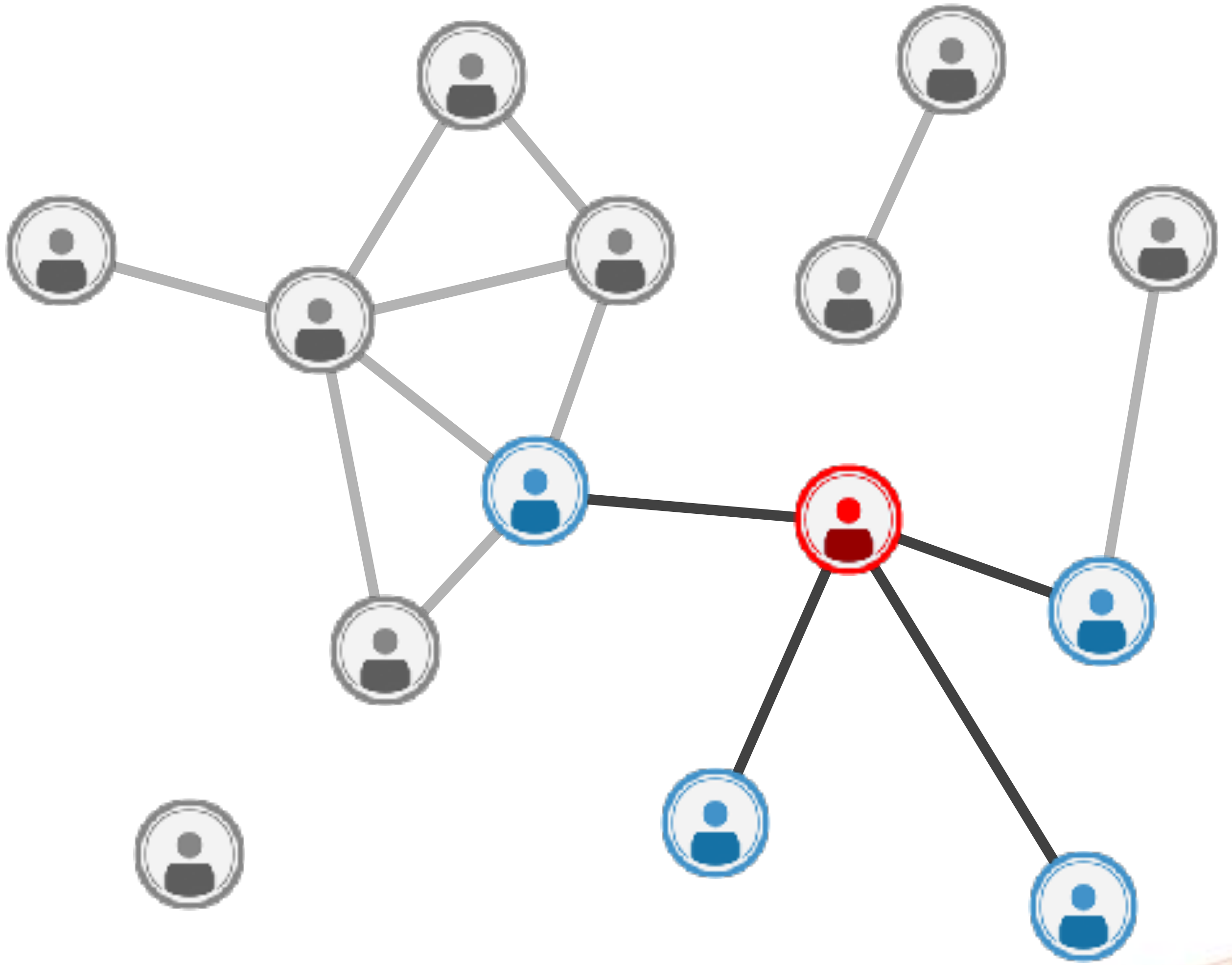


VIS 2017

Complex Networks

A network can be represented by a **graph** ***G=(V, E)*** such that:

individuals	⟺	***nodes***
ties		***edges***
tie strength		***edge weight***



Network Analysis:

- *Retrieval of **global** structure*
- ***Local** study of each node*

VIS 2017

Main Goal

Propose a **general method** for

- performing ***global and local analyses***
- ***interactively visualizing*** network structure according to different parameters in a single view
- ***comparing*** and characterize ***different networks***

Key Idea:

Adopt a *persistence-like* approach to *clique community* decomposition

Outline

- Clique Communities
- Clique Community Persistence
- Single Network Analysis
  - *Interactive Visualization Tool based on Nested Graphs*
- Network Comparison
  - *Distance based on Persistence Indicator Functions*
  - *Clique Community Centrality Measure*
- Conclusions and Future Developments

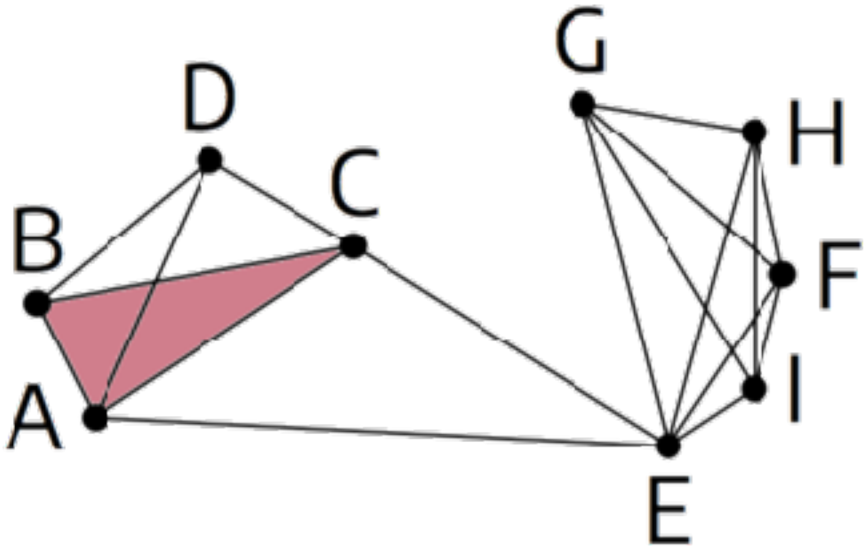
VIS 2017

Clique Communities

Given a network $G=(V, E)$,

***k*-clique:**

A **complete subgraph** of ***k*** vertices of G



3-clique

[Palla et al., 2007]

Clique Communities

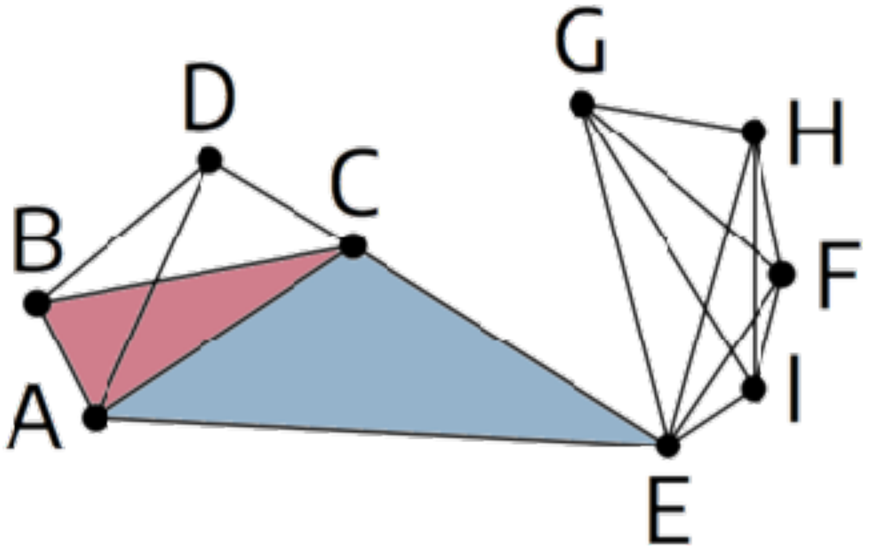
Given a network *G=(V, E)*,

k-clique:

A **complete subgraph** of *k* vertices of *G*

k-adjacency:

Two *k*-cliques are ***k*-adjacent** if they **share *k-1* nodes**



3-adjacent

VIS 2017

[Palla et al., 2007]

Clique Communities

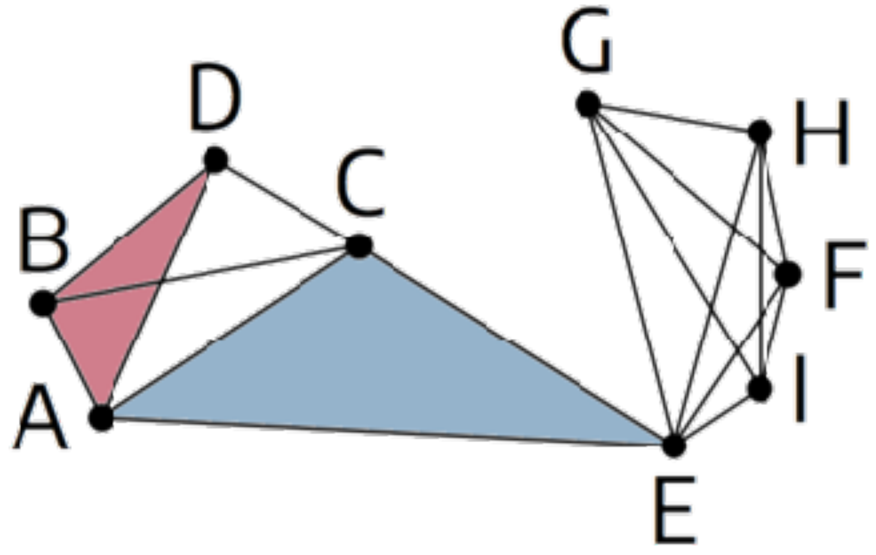
Given a network ***G=(V, E)***,

k-clique:

A **complete subgraph** of ***k*** vertices of *G*

k-adjacency:

Two *k*-cliques are ***k*-adjacent** if they **share *k-1* nodes**



non 3-adjacent

VIS 2017

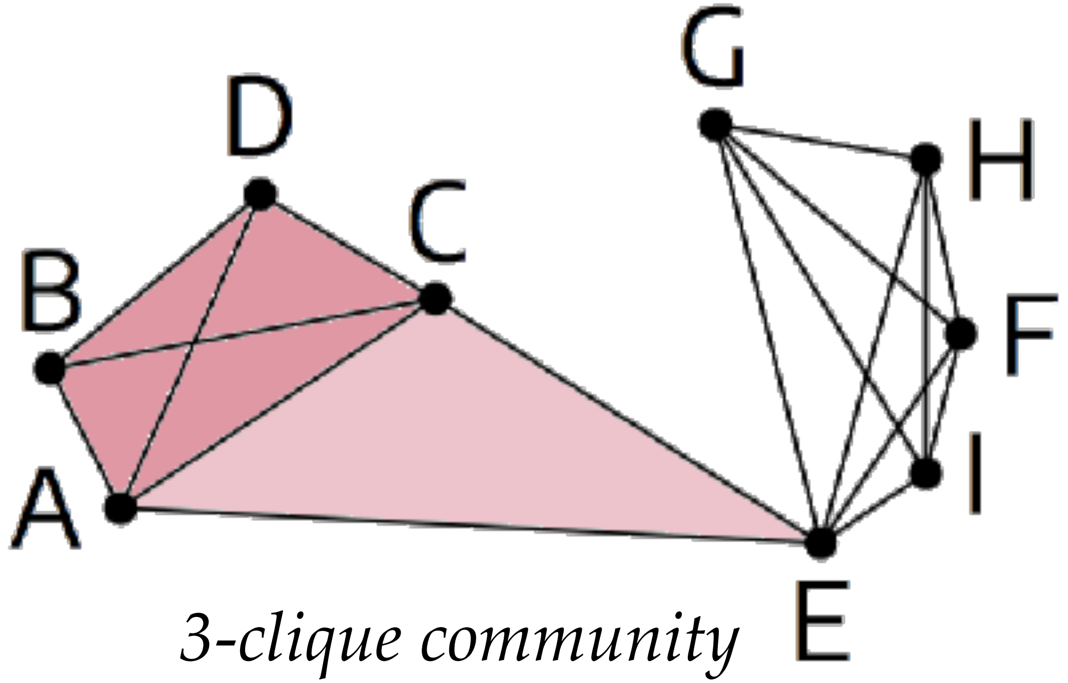
[Palla et al., 2007]

Clique Communities

Given a network *G=(V, E)*,

***k*-clique Community:**

Maximal union of *k*-cliques ***pairwise connected***
by a ***sequence of k-adjacent cliques***



3-clique community

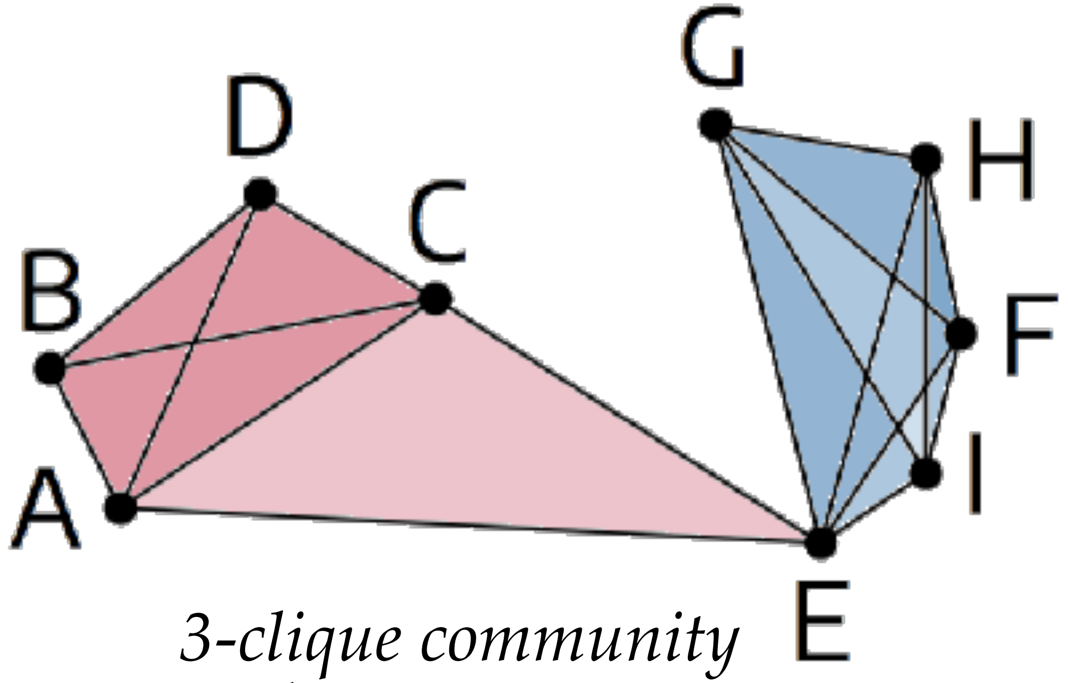
[Palla et al., 2007]

Clique Communities

Given a network ***G=(V, E)***,

k-clique Community:

Maximal union of k-cliques ***pairwise connected*** by a ***sequence of k-adjacent cliques***



3-clique community decomposition

k-clique Community Decomposition:

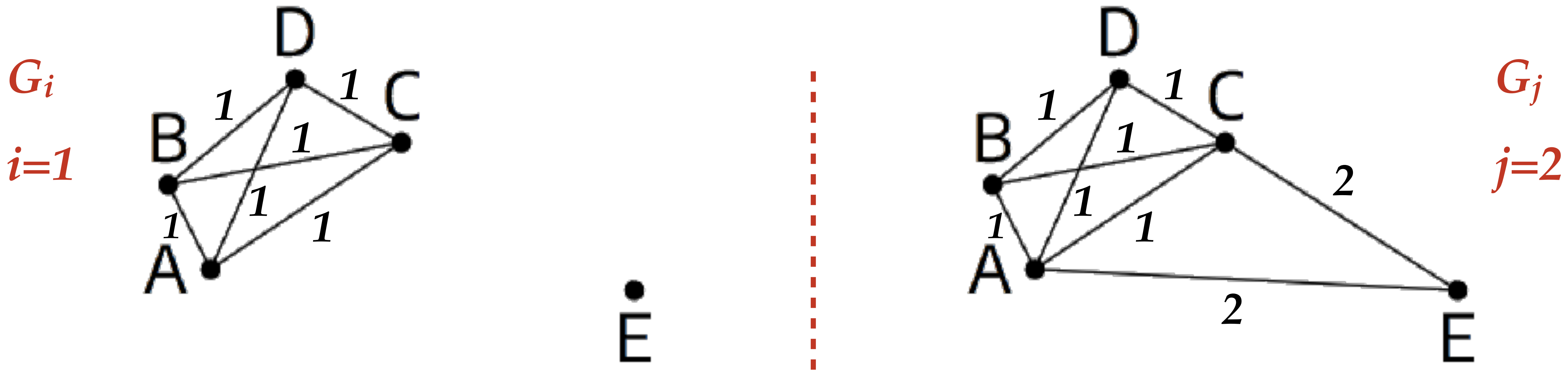
The **partition** of the k-cliques of G ***induced by the k-clique communities***

[Palla et al., 2007]

Clique Community Persistence

Clique Communities & Weighted Networks:

Given a weighted network G and two threshold values $i < j$,

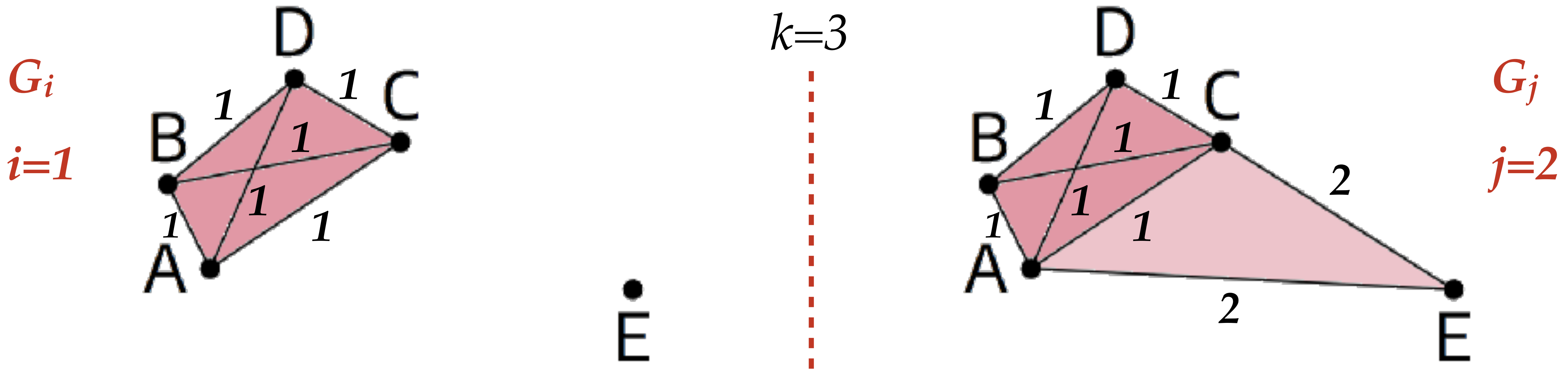


G_i is contained in G_j

Clique Community Persistence

Clique Communities & Weighted Networks:

Given a weighted network ***G*** and two threshold values ***i*** < ***j***,



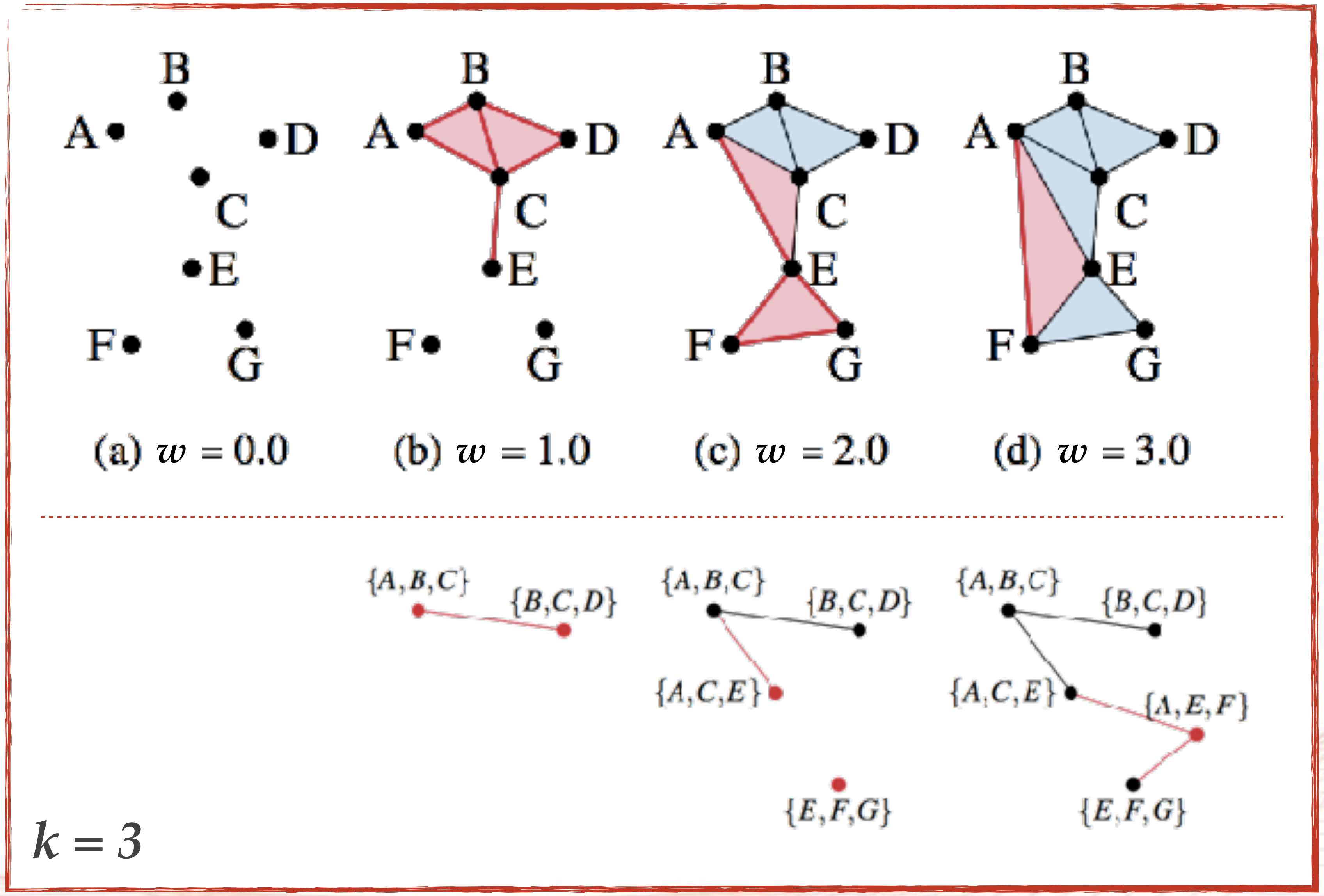
Each k-clique community of G_i is **contained** in ***exactly one***
k-clique community of G_j

Clique Community Persistence

Fixing a value for *k* and varying the edge-weight threshold:

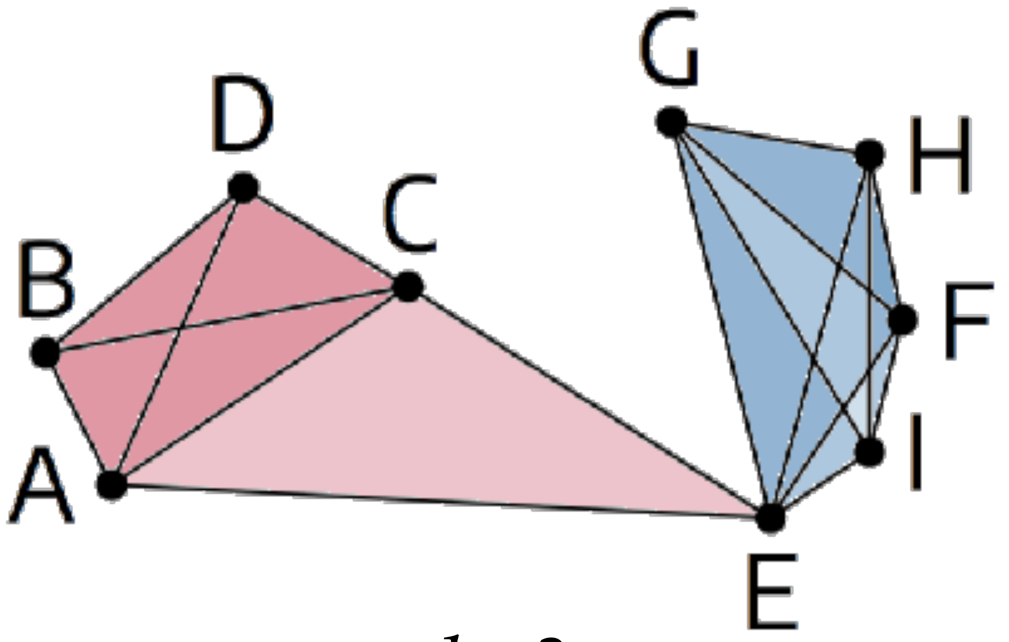
The evolution of k-clique communities of G can be tracked by:

- building a sequence of *k*-dual graphs:
  - *vertices ↔ k-cliques*
  - *edges ↔ adjacent k-cliques*
- tracking the **connected components** of the sequence of *k*-dual graphs

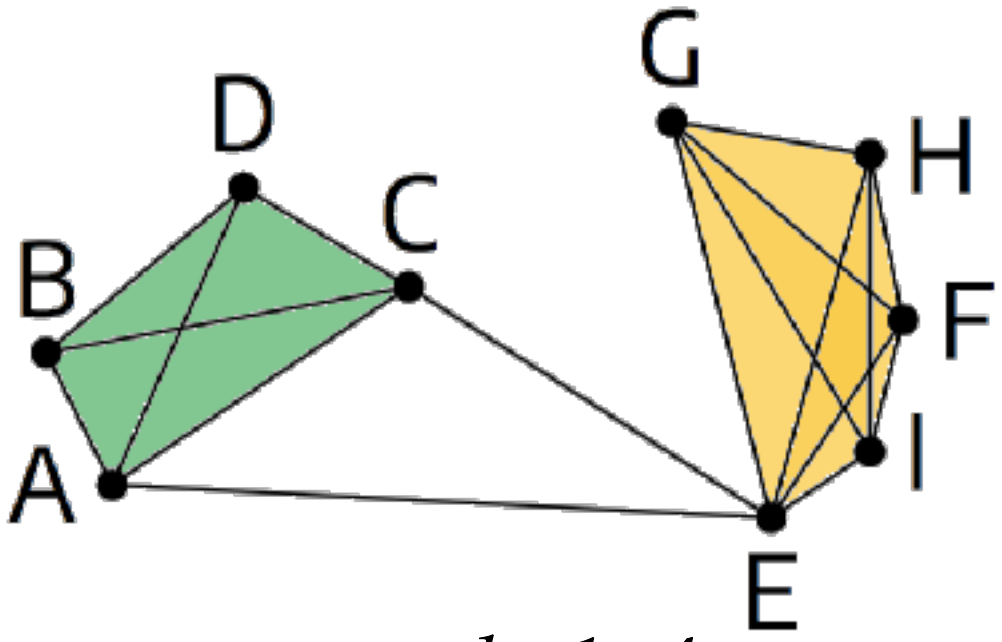


Clique Community Persistence

Clique Communities & Multiple *k* Values:



k=3



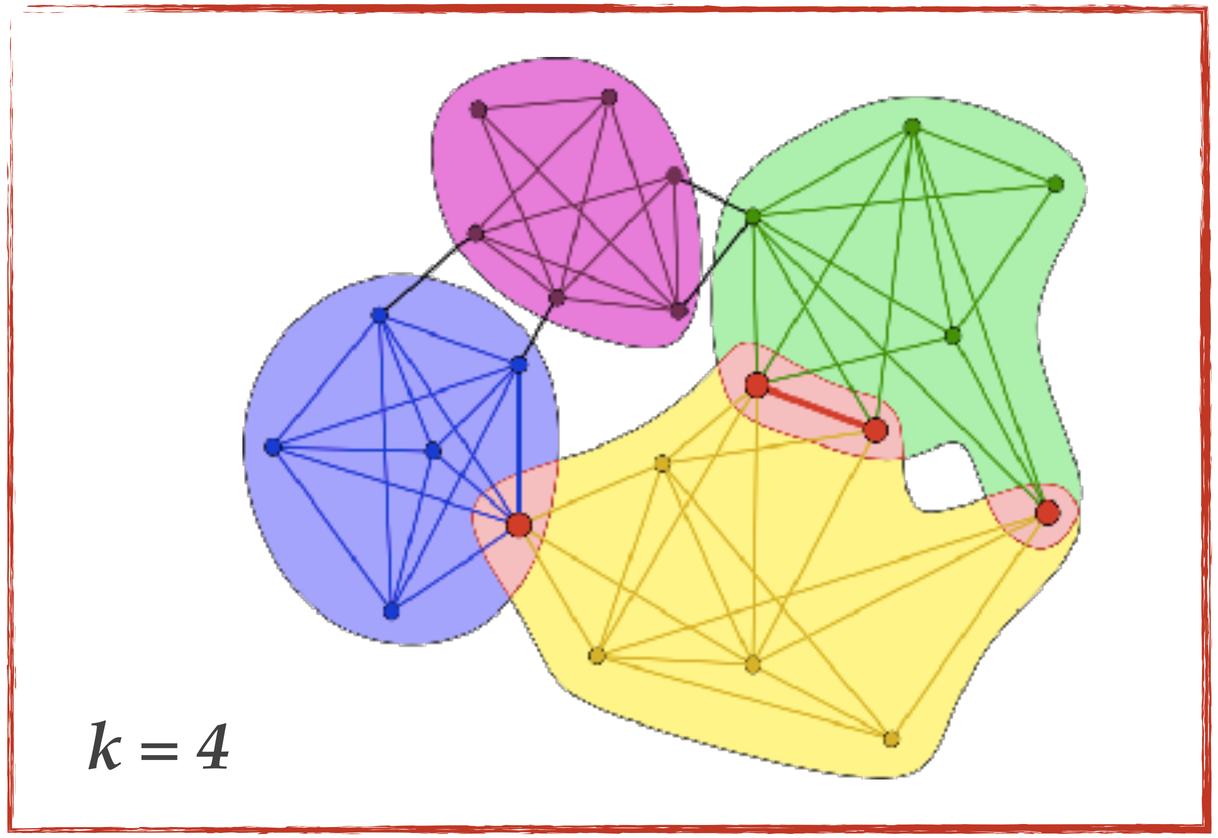
k+1=4

Each *(k+1)*-clique community of *G* is **contained** in ***exactly one*** *k*-clique community of *G*

Clique Community Persistence

Pros & Cons:

- *Reveal Highly Connected Communities*
- *Allow Overlaps*
- *Display a (double) Hierarchical Structure*



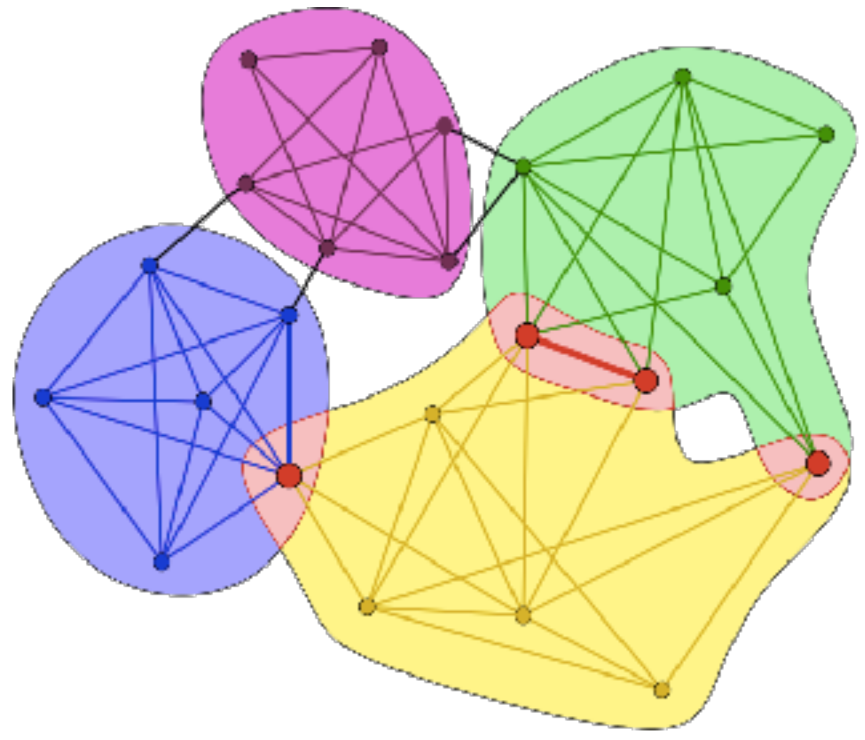
Focusing on a single value for k and weight threshold w provides just a partial view of the network structure

VIS 2017

Clique Community Persistence

Pros & Cons:

- *Reveal **Highly Connected** Communities*
- *Allow **Overlaps***
- *Display a (double) **Hierarchical Structure***



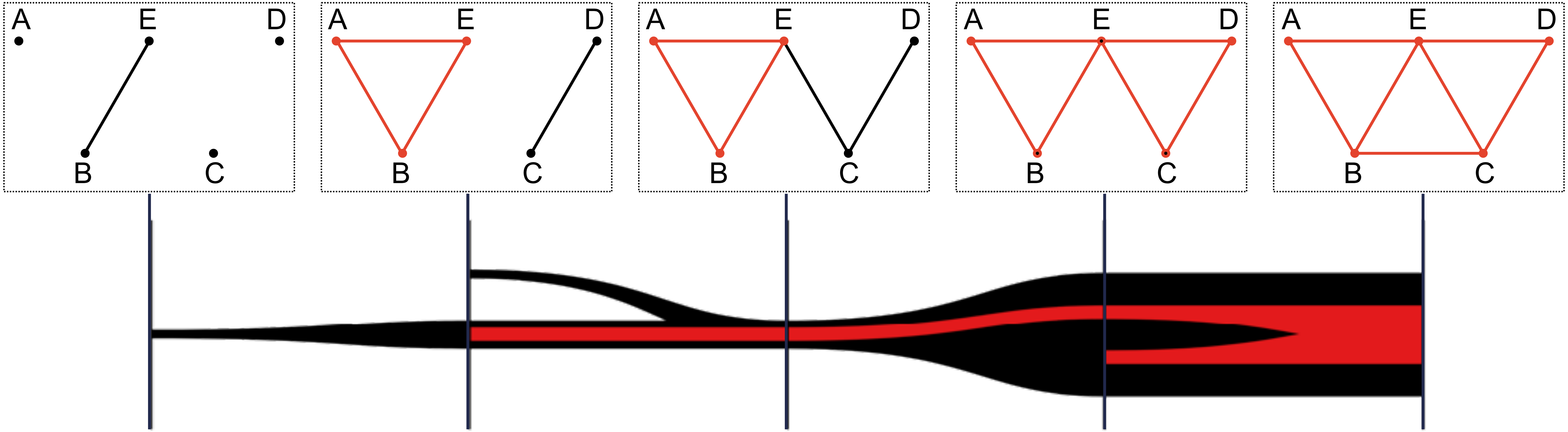
$k = 4$

Develop a tool for simultaneously dealing with multiple k and w values

Clique Community Persistence

Nested Graph: *[Lukasczyk et al., EuroVis 2017]*

- Originally defined for connected components in scalar fields
- Illustrates ***evolutions across two parameters***

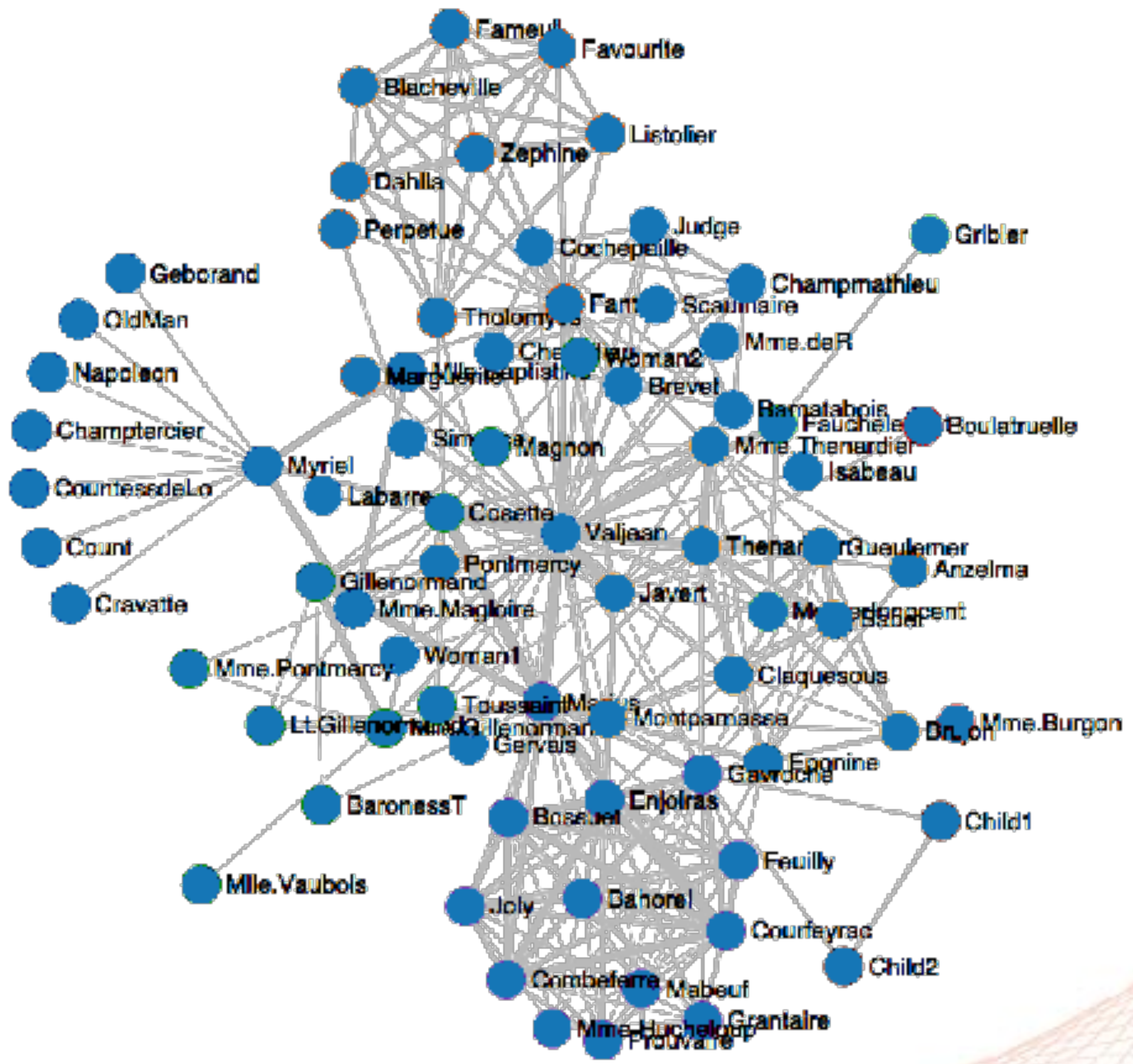


Single Network Analysis

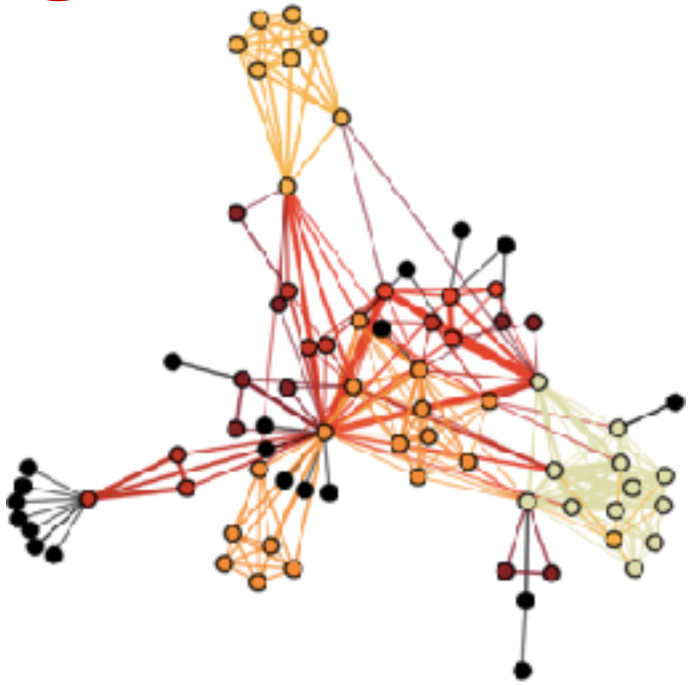
"Les Misérables" Network:

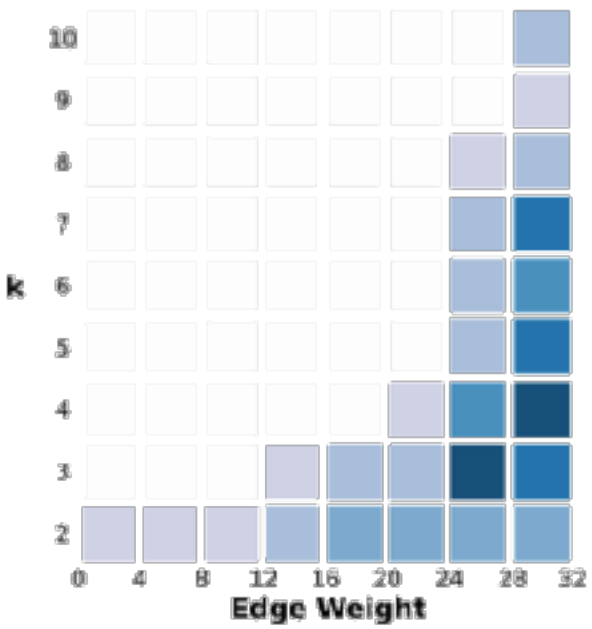
- ***Co-occurrence network*** between the characters in **Victor Hugo's novel** *"Les Misérables"*
  - *77 nodes*
  - *254 edges*

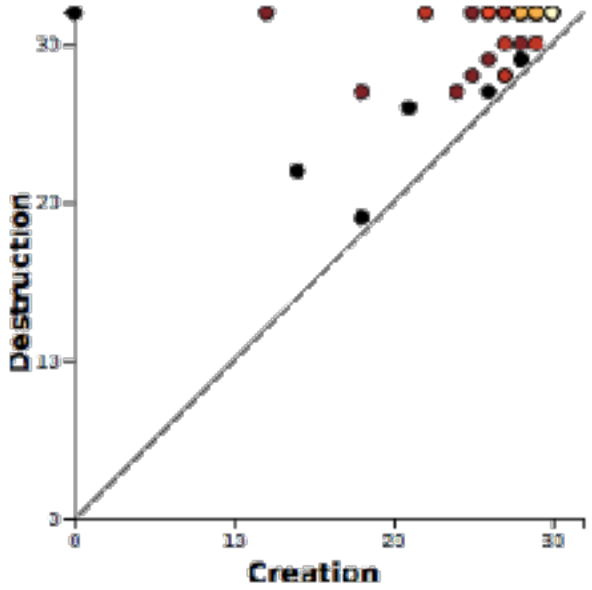
- ***edge weight*** ↔ ***number of co-occurrences*** *between two characters* (*inverse of the*)

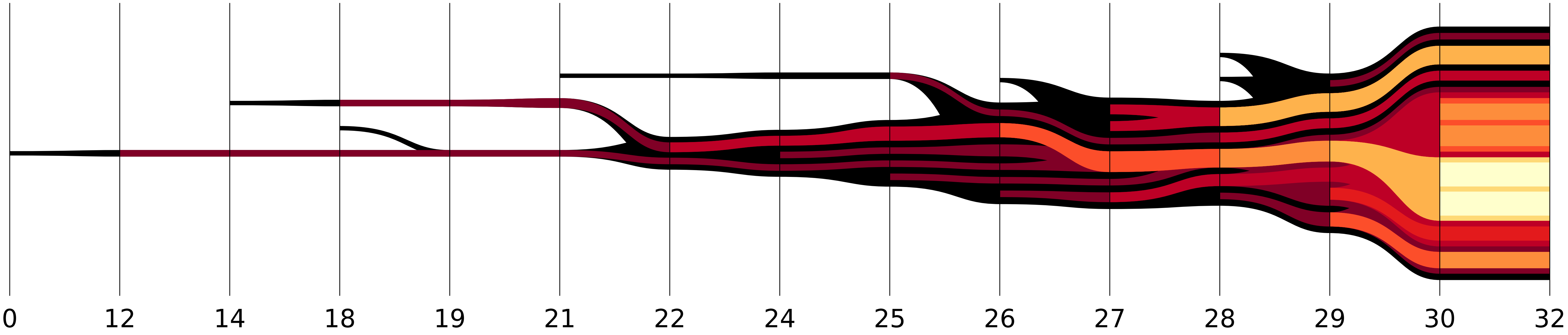


Single Network Analysis



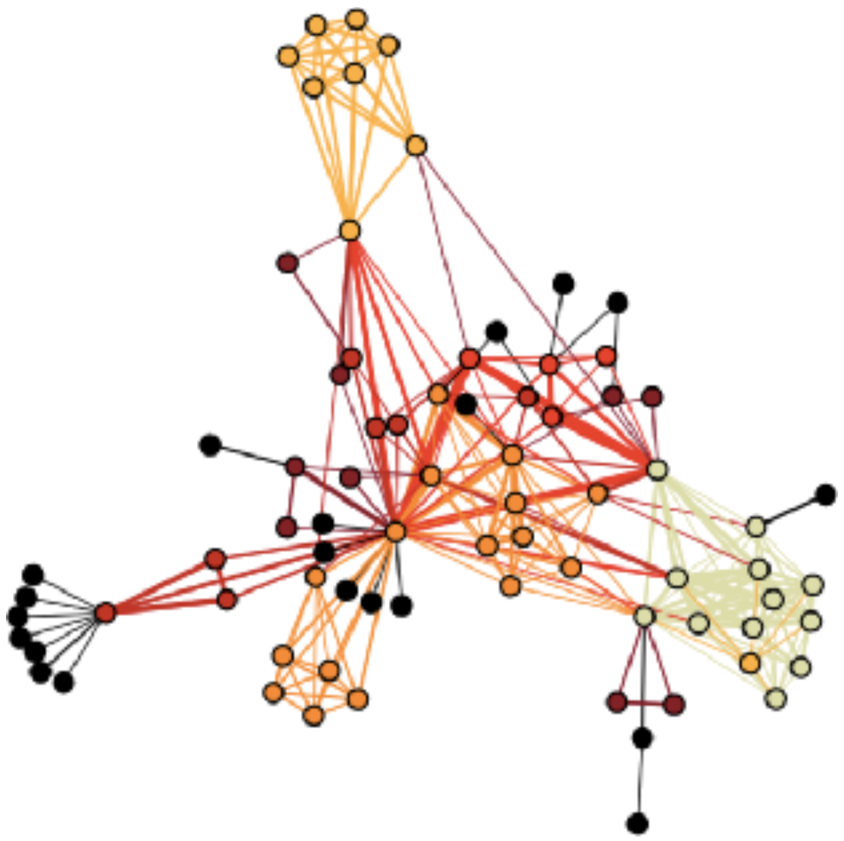


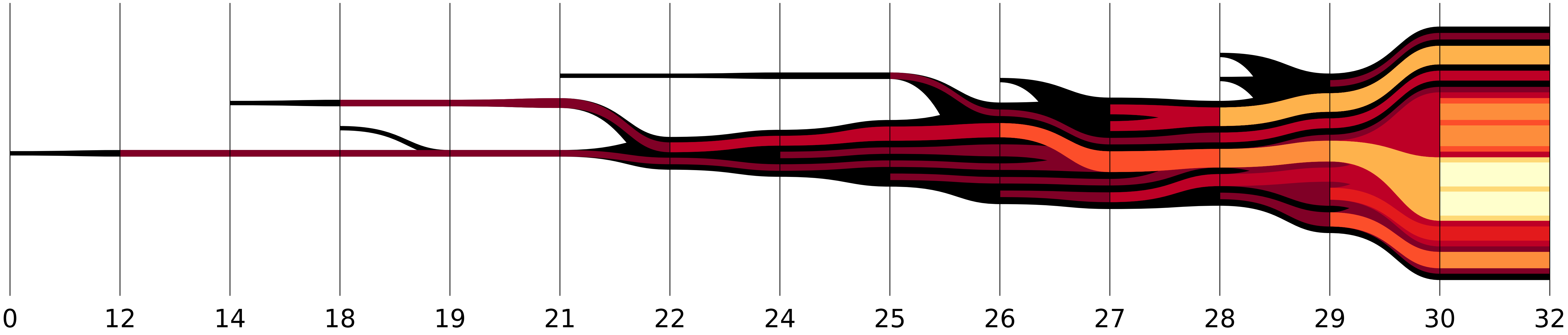




Single Network Analysis

Nested-based visualization tool allows the user for

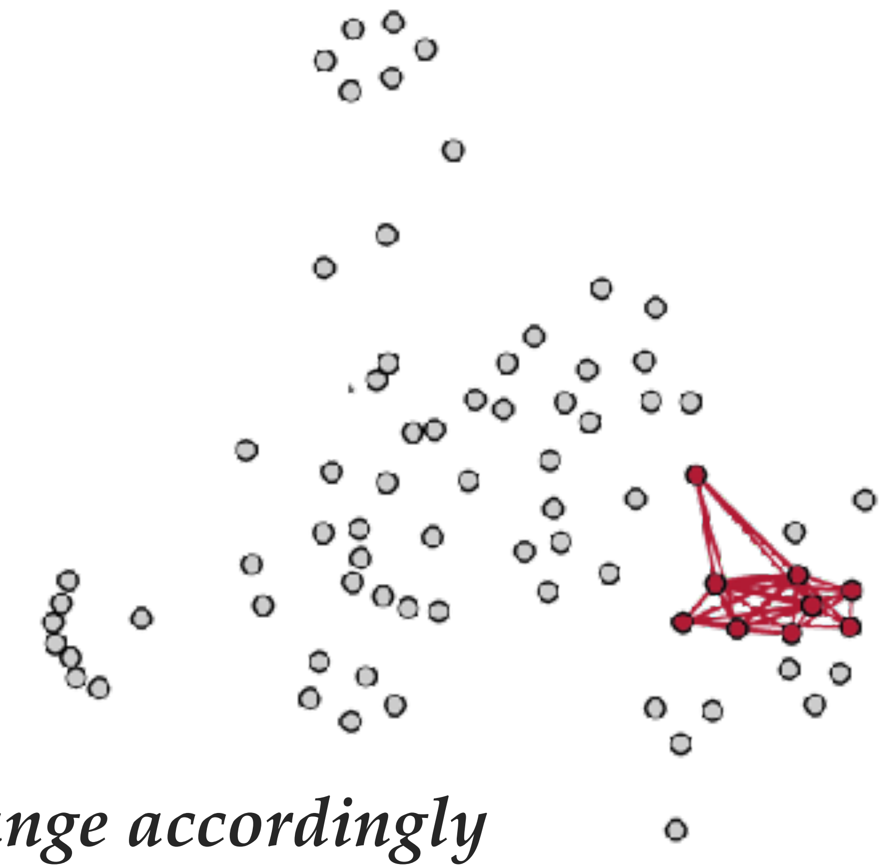




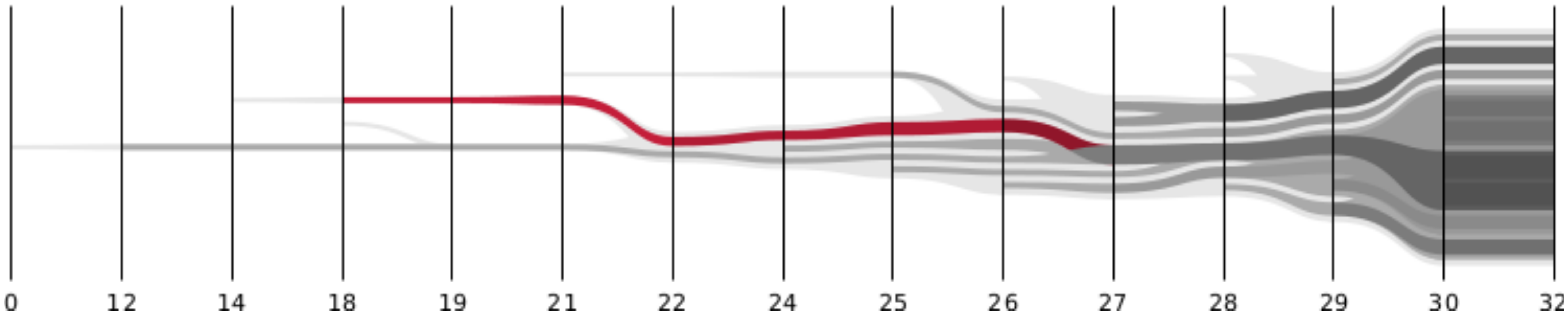
Single Network Analysis

Nested-based visualization tool allows the user for

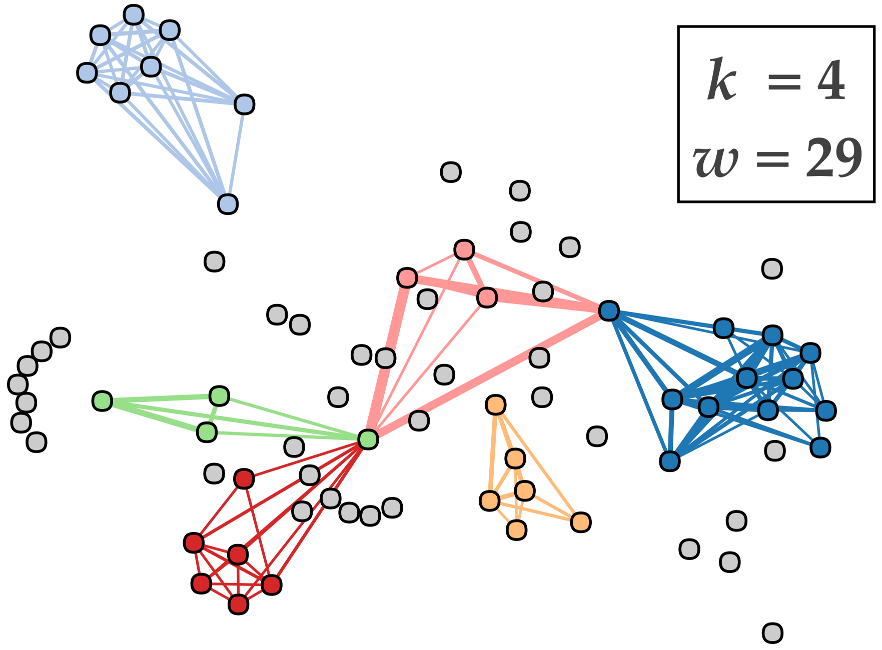
- ***focusing on*** *the evolution of a* ***specific clique community***
- ***selecting*** *and* ***interactivity exploring different edge weights*** *and* ***clique degrees***



while the force-directed graph layout and the nested graph ***change accordingly***



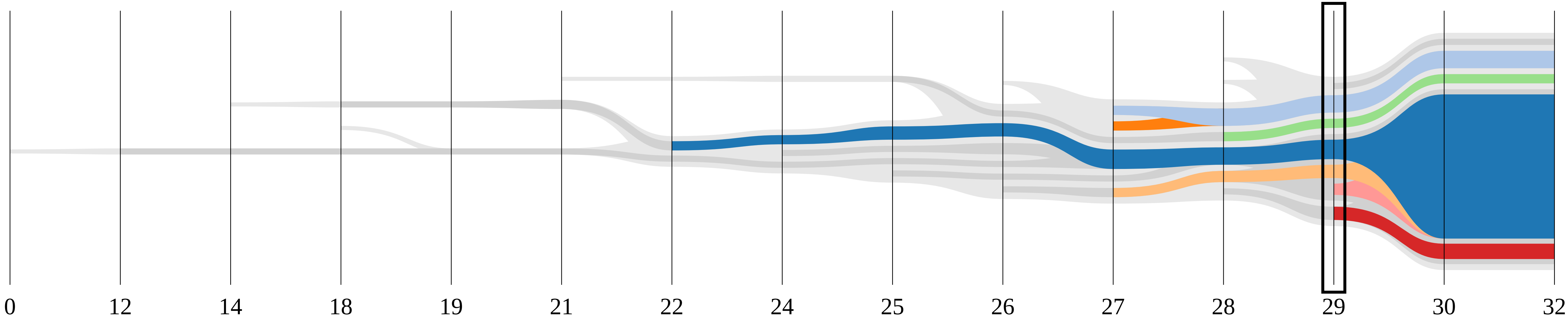
Single Network Analysis



Nested-based visualization tool allows the user for

- *focusing on the evolution of a specific clique community*
- ***selecting** and **interactivity exploring different edge weights** and **clique degrees***

while the force-directed graph layout and the nested graph ***change accordingly***



Single Network Analysis

Nested-based visualization tool allows the user for

- ***focusing on*** *the evolution of a* ***specific clique community***
- ***selecting*** *and* ***interactivity exploring different edge weights*** *and* ***clique degrees***

while the force-directed graph layout and the nested graph ***change accordingly***

Intuitively:

edge-weight *variation* ↔ *reveal the* ***core part of a communitiy***

clique-degree *variation* ↔ *analyze* ***community according to different granularities***

Network Comparison

Clique community persistence enables the introduction of ***comparison measures***:

- *Persistence indicator function (PIF)*
- *PIF-based distance*
- *Clique community centrality measure*

Network Comparison

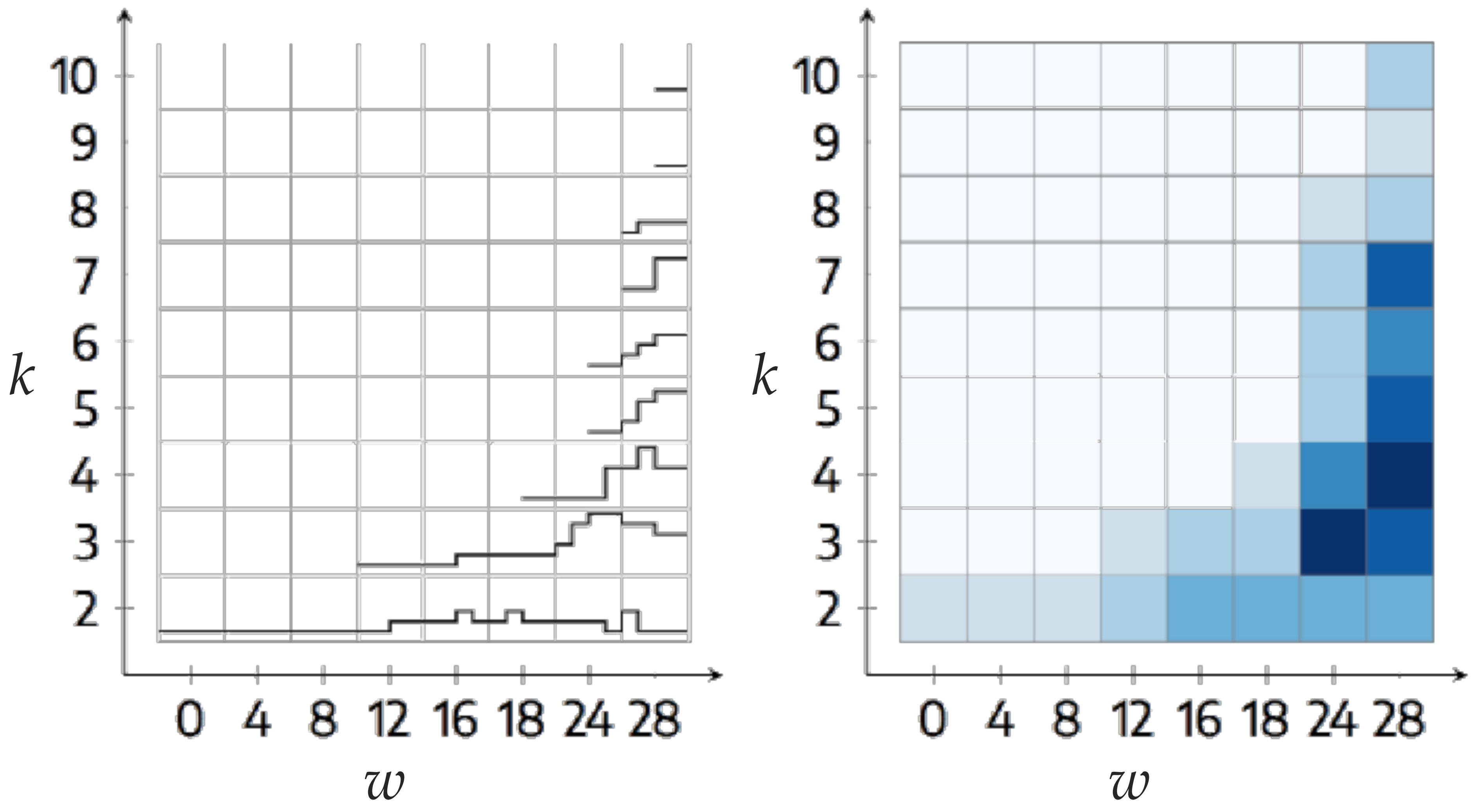
Persistence Indicator Function:

Defined as the function

$$f_k : \mathbb{R} \longrightarrow \mathbb{N}$$

assigning

$$w \longmapsto \begin{array}{c} \textit{\# k-cliques} \\ \textit{communities "alive"} \\ \textit{at threshold } w \end{array}$$



VIS 2017

Network Comparison

PIF-based Distance:

Given two persistence indicator functions *f* and *g*,

PIF-based distance is defined to be the L_p distance between *f* and *g*:

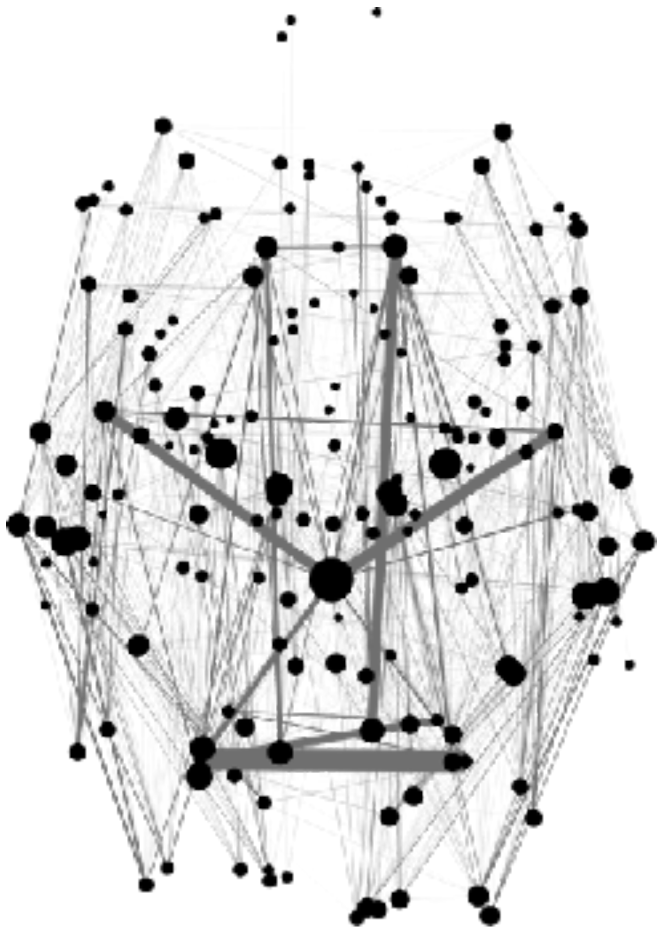
$$dist(f,g) = \left(\int_{\mathbb{R}} |f(x) - g(x)|^p \, dx \right)^{\frac{1}{p}}$$

- Quantifies dissimilarities between PIFs
- Easier to be computed than Wasserstein and bottleneck distances
- Highly correlated to Wasserstein distance

Network Comparison

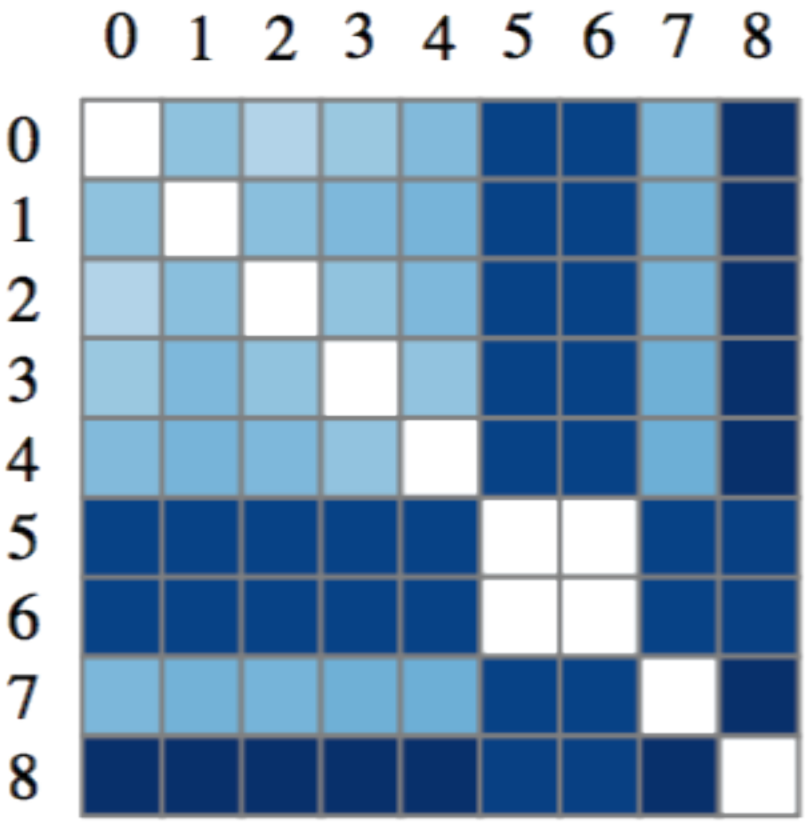
Brain Networks:

- ***Biological networks*** representing variants of ***human brain connectivity***
  - *9 instances considered*
- ***nodes*** ↔ ***brain areas***
- ***edges*** ↔ ***fibers connecting different areas***



Network Comparison

Brain Networks:



Variant	Density	Diam. (weighted)	Avg. degree (weighted)
0	0.125	4 (60.0)	21.21 (2300.3)
1	0.124	4 (60.0)	21.06 (2296.0)
2	0.124	4 (60.0)	21.13 (2295.2)
3	0.124	4 (60.0)	21.16 (2282.0)
4	0.124	4 (60.0)	21.15 (2279.3)
5	0.125	4 (60.0)	21.19 (2264.0)
6	0.125	4 (60.0)	21.19 (2264.0)
7	0.124	4 (60.0)	21.16 (2279.6)
8	0.125	4 (60.0)	21.20 (2257.5)

PIF-based distance reveals differences between networks that common graph measures are incapable of detecting

Network Comparison

Clique Community Centrality:

Clique community centrality of a node v is defined as

$$centrality(v) = \sum_{C \ni v} pers(C)$$

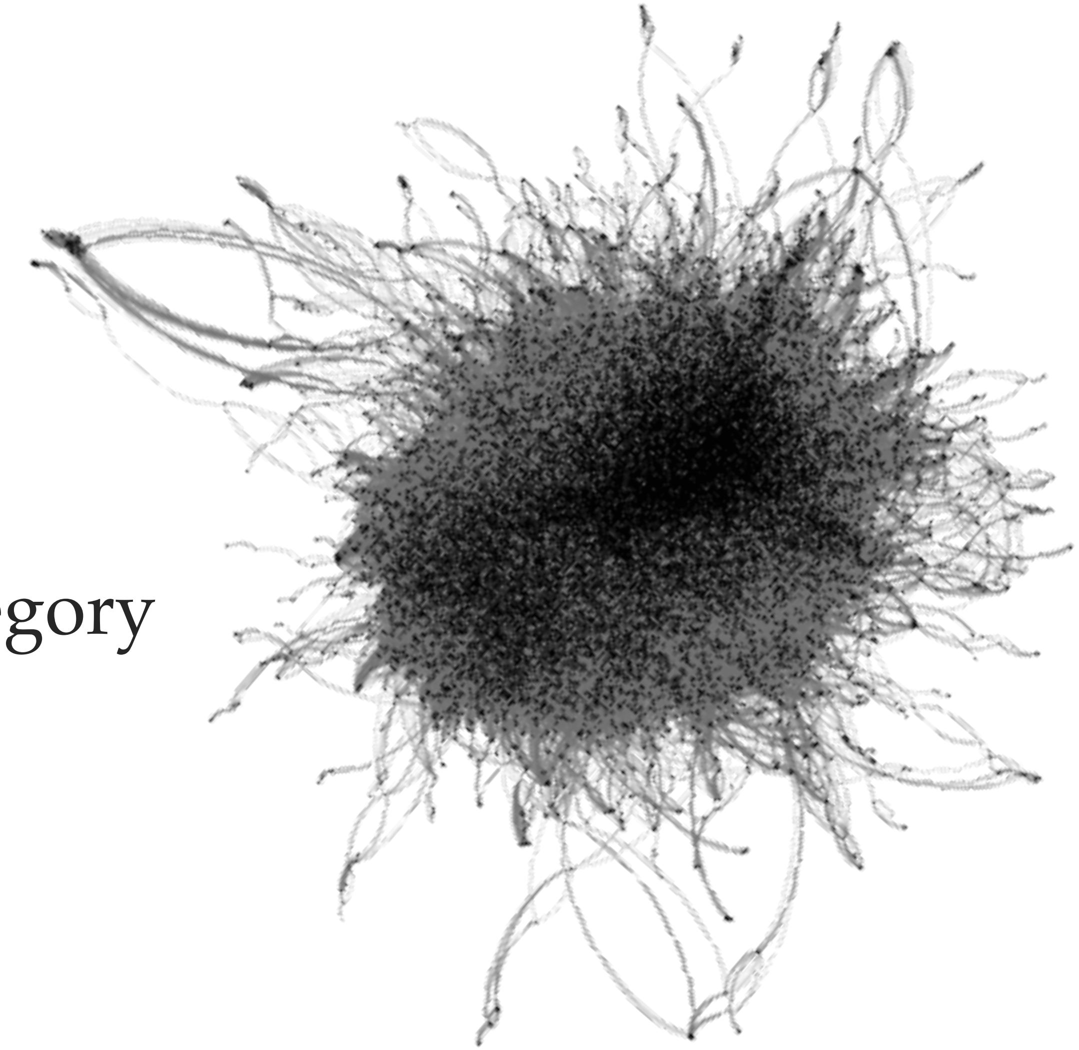
where

- C is any clique community containing v
- ***pers(C)*** is the "lifespan" of C

*Nodes belonging to **high-persistence communities** are identified as **relevant***

Network Comparison

Condensed matter collaboration:



- ***Collaborative networks*** describing ***scientist co-authorship of the "Condensed Matter"*** arXiv category
  - ***3** snapshots in time considered (1999, 2003, 2005)*
- Network sizes:
  - ***16K - 40K** nodes*
  - ***47K - 175K** edges*

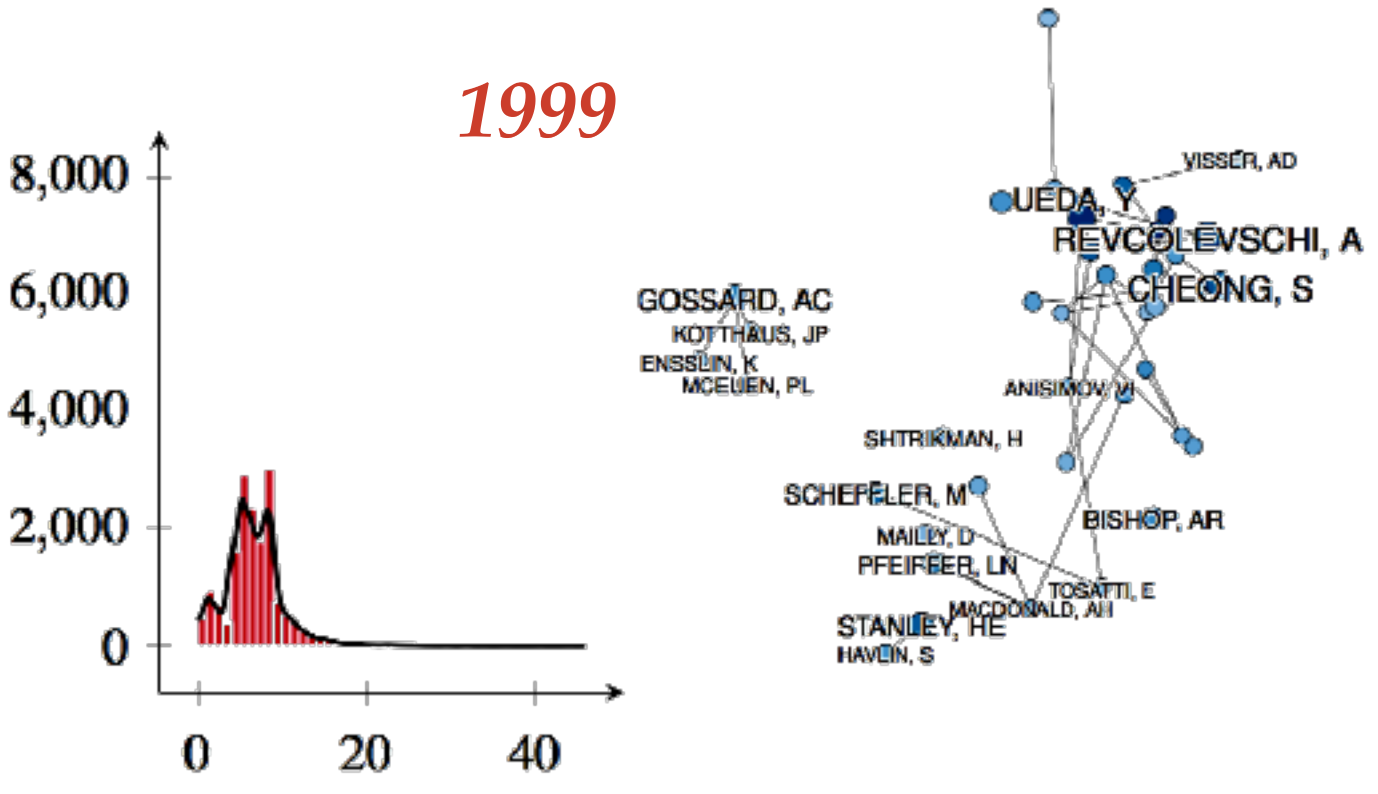
Clique community centrality *allows for*

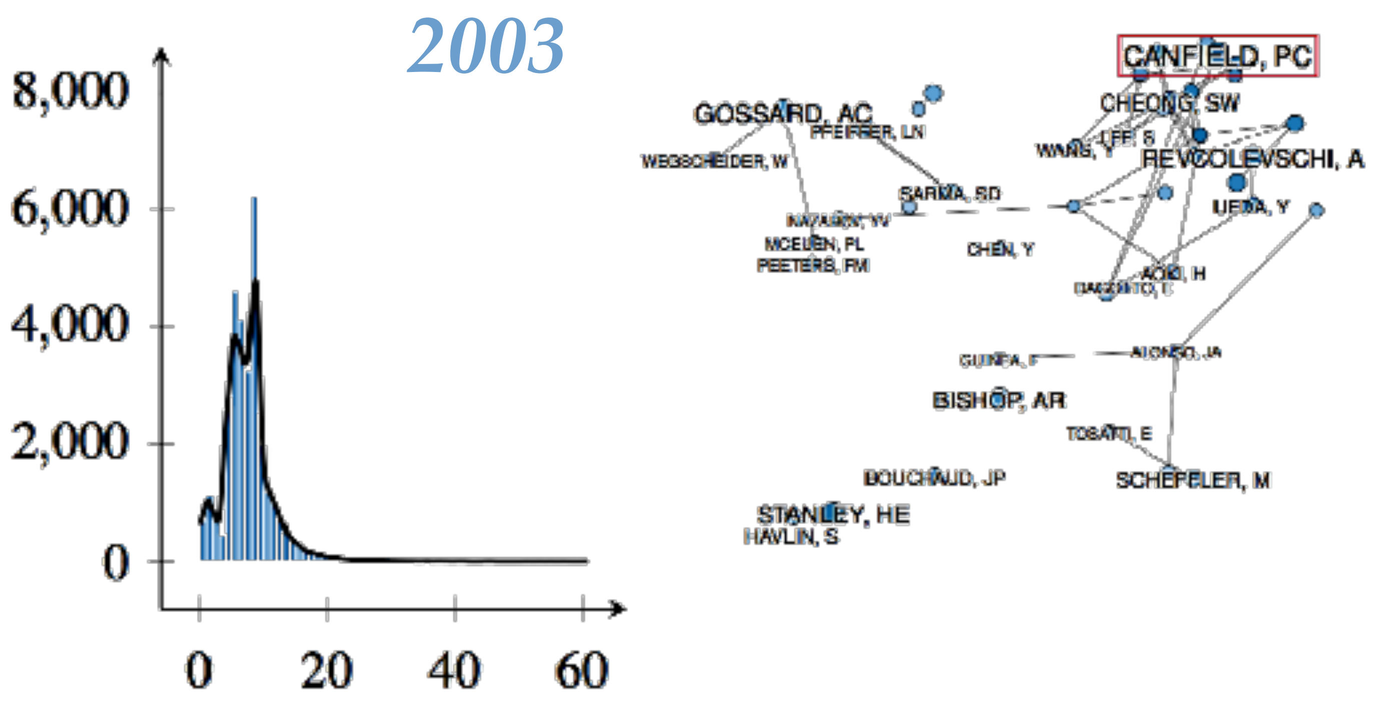
- *evaluating the **evolution of network connectivity***
- *filtering away the **less relevant nodes***

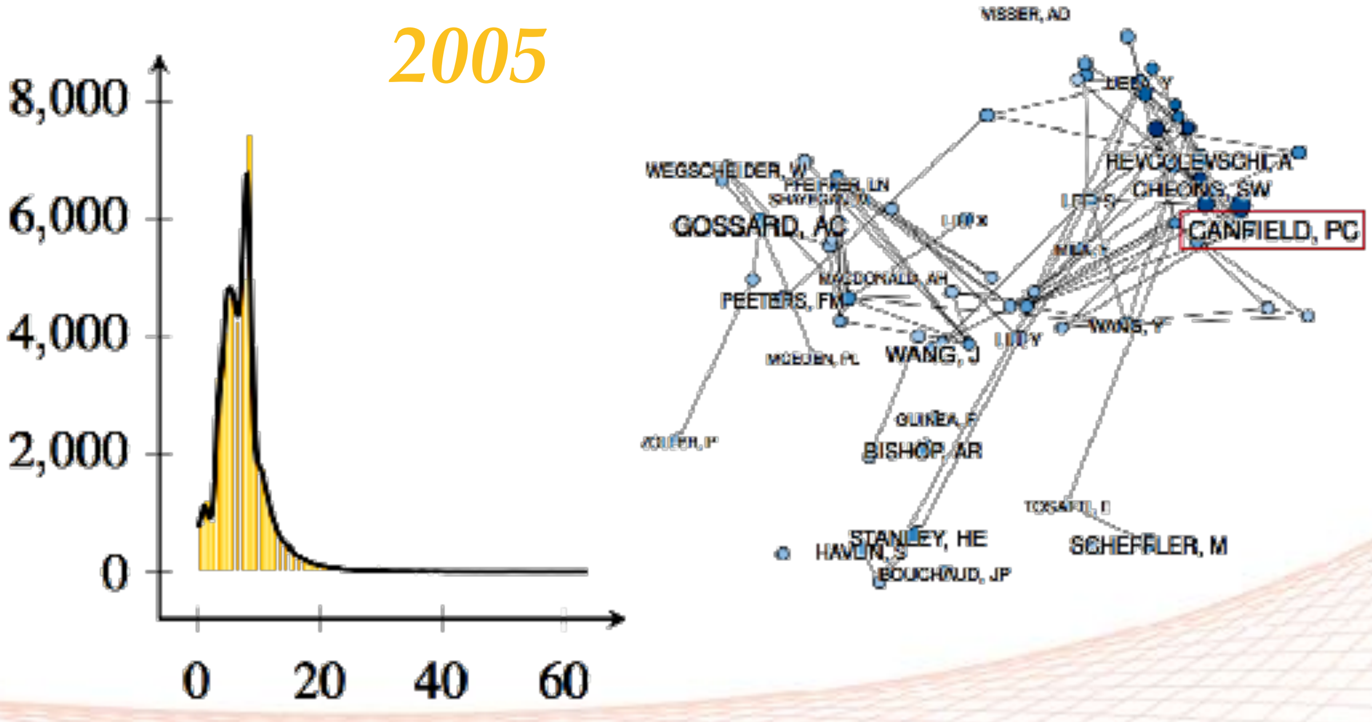
VIS 2017

Network Comparison

Condensed matter collaboration:







Density estimates of the clique community centrality values

Conclusions

To Summarize:

In our work, we propose a ***new method*** based on ***clique community persistence*** for analyzing ***global*** and ***local properties*** of complex networks

This method leads to the design and the introduction of

- *an **interactive visualization tool** based on nested graphs*
- *new criteria and distances for **network comparison***

Future Developments:

- *Extend to **time-varying** non-weighted networks*
- *Improve **clique community computation***

VIS 2017

Thank you

Ulderico Fugacci

TU Kaiserslautern, *Dept. of Computer Science*