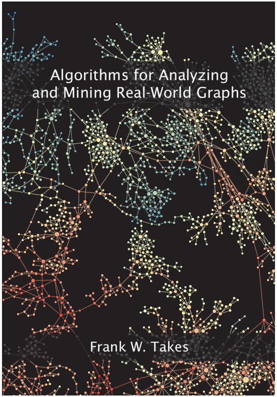
Discovering the Building Blocks of Social Networks

Frank Takes

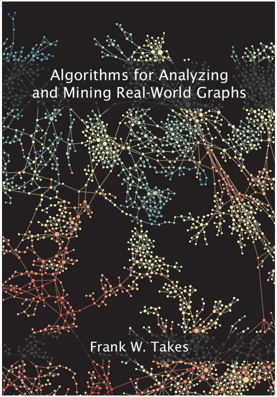
Leiden University

NETWORKS @ Kaap Doorn
January 20, 2020

Introduction

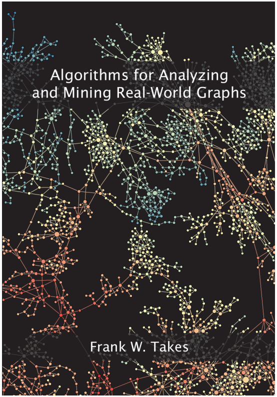


Introduction



graph algorithms, network science, complex networks,

Introduction



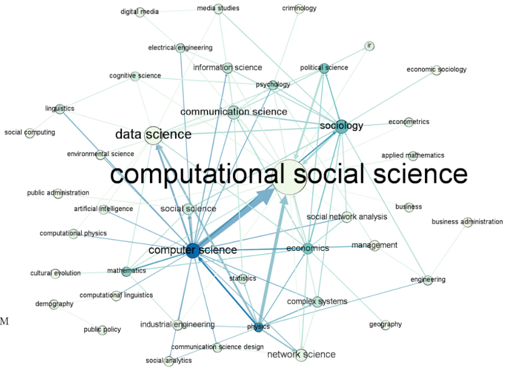
graph algorithms, network science, complex networks,
social network analysis, computational social science

Computational social science

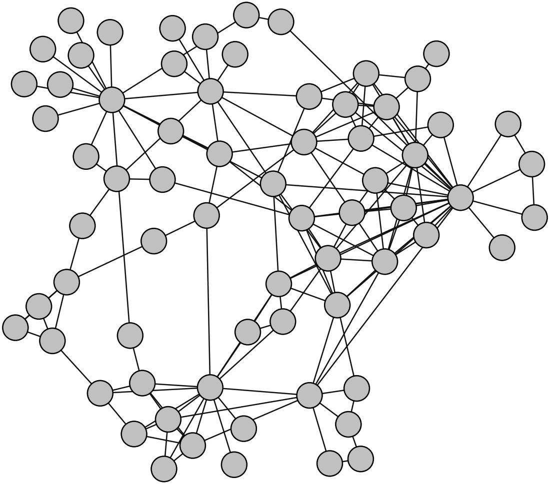
5th International Conference on Computational Social Science IC2S2 2019

https://ic2s2.org

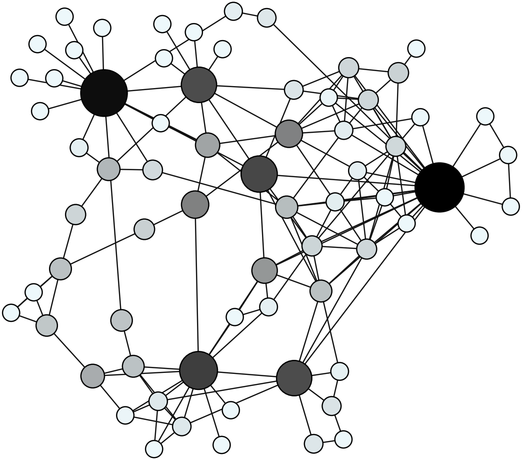
UNIVERSITY OF AMSTERDAM



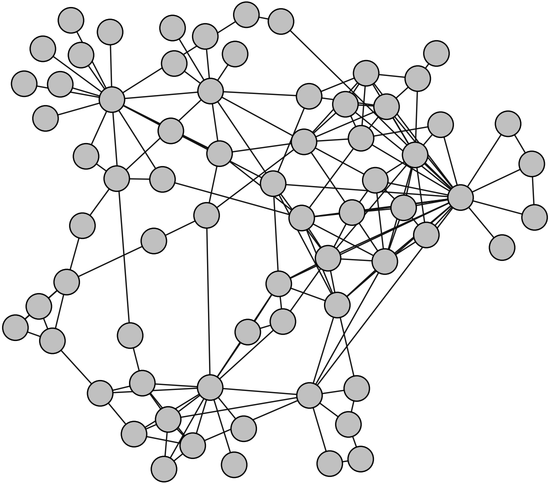
Network analysis



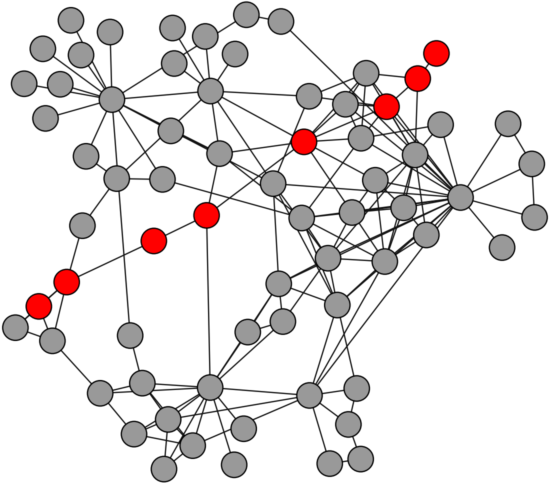
Micro scale



Macro scale



Macro scale



Network analysis

- **Micro** scale: analyzing the position of individual nodes, based on their structural position in the network (e.g., node centrality, etc.)
- **Macro** scale: analyzing the structure of the network as a whole (e.g., network diameter, small-world effect, etc.)

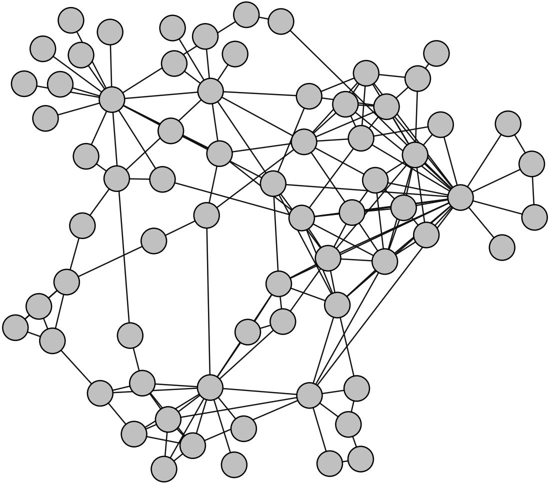
Network analysis

- **Micro** scale: analyzing the position of individual nodes, based on their structural position in the network (e.g., node centrality, etc.)
- **Macro** scale: analyzing the structure of the network as a whole (e.g., network diameter, small-world effect, etc.)
- **Meso** scale: analyzing groups of nodes occurring in a particular configuration

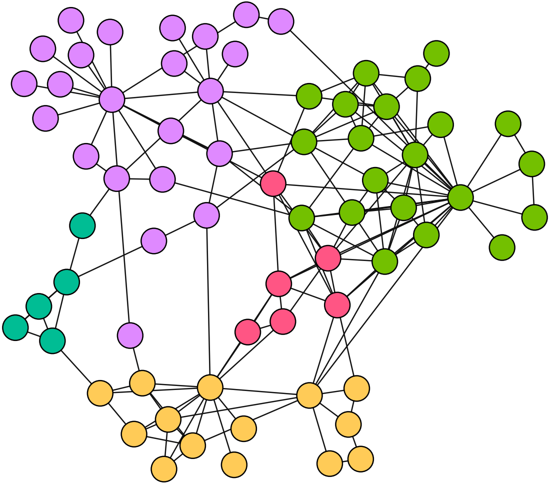
Network analysis

- **Micro** scale: analyzing the position of individual nodes, based on their structural position in the network (e.g., node centrality, etc.)
- **Macro** scale: analyzing the structure of the network as a whole (e.g., network diameter, small-world effect, etc.)
- **Meso** scale: analyzing groups of nodes occurring in a particular configuration (e.g., communities or networks motifs)

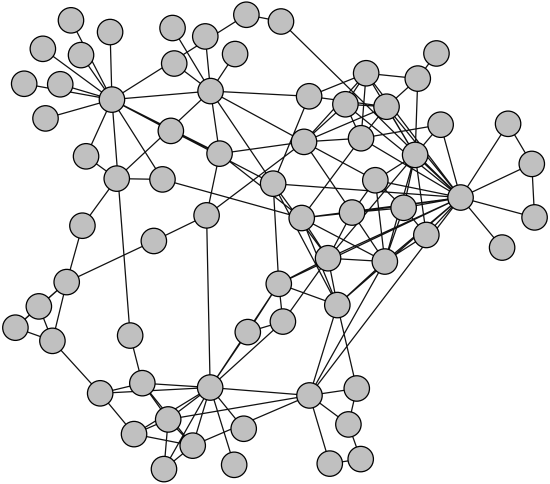
Meso scale: communities



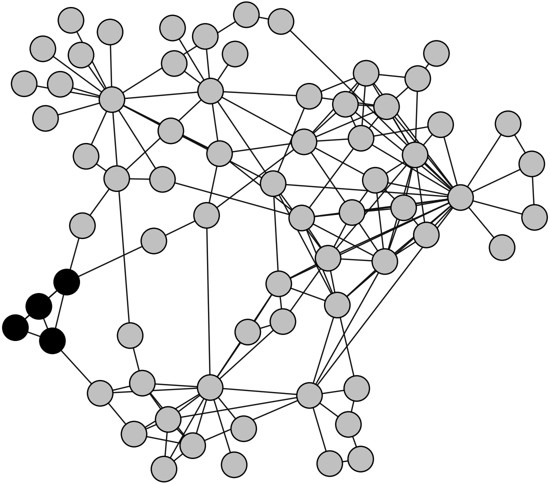
Meso scale: communities



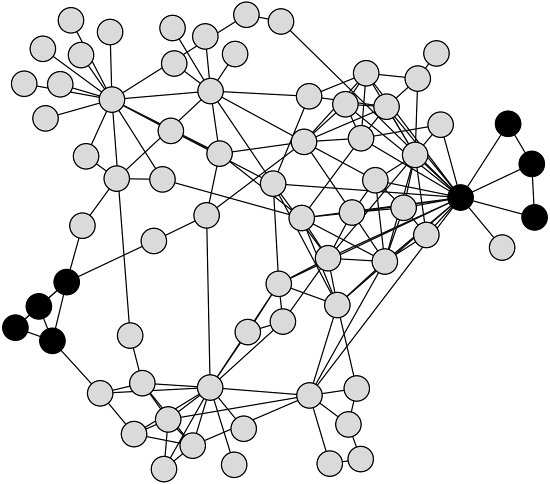
Meso scale: motifs



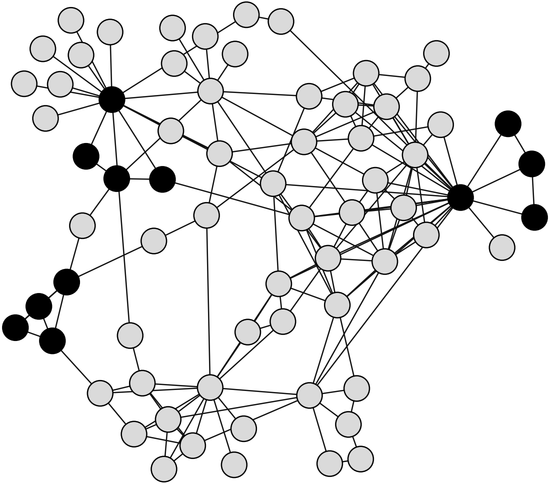
Meso scale: motifs

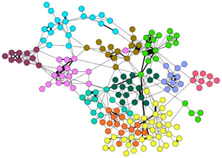


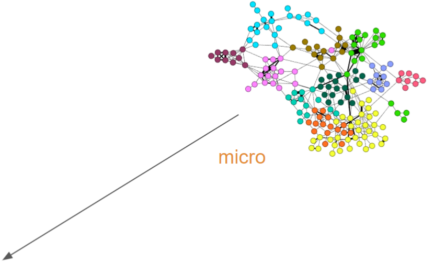
Meso scale: motifs

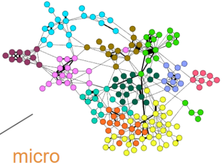


Meso scale: motifs



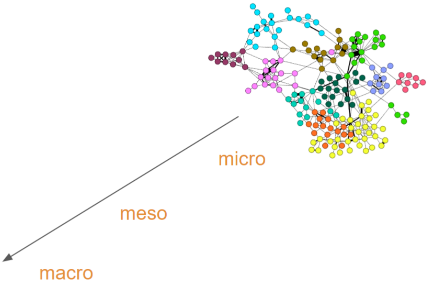


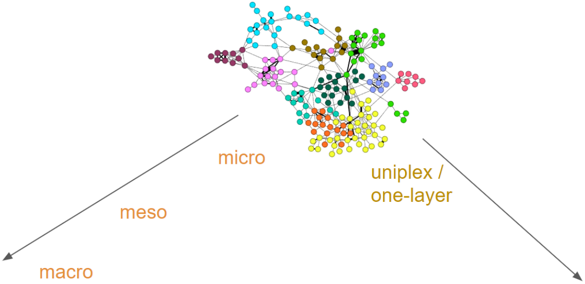

micro

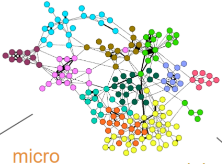


micro

macro


micro
meso
macro


micro
meso
macro
uniplex /
one-layer



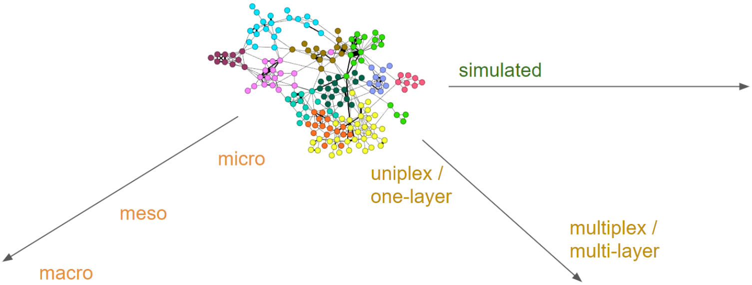
micro

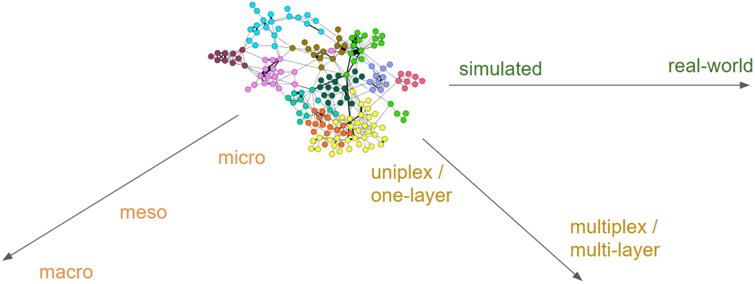
meso

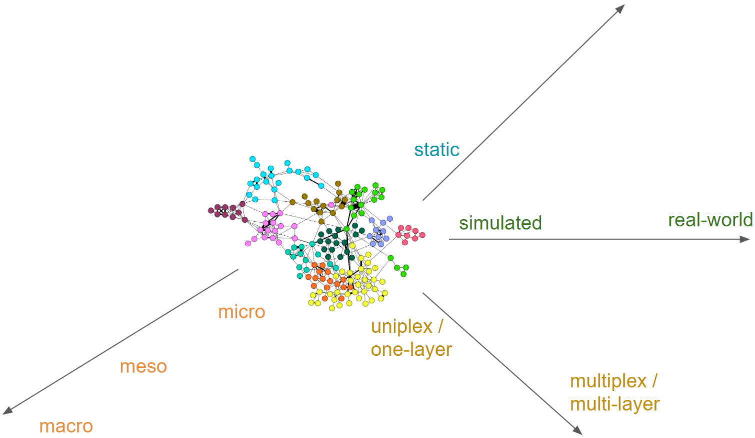
macro

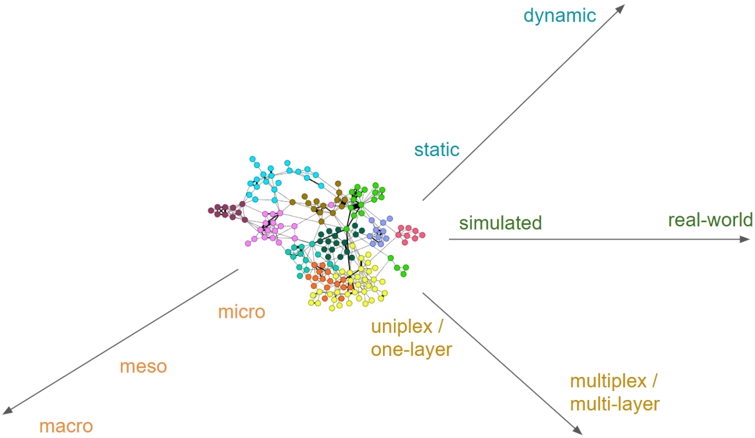
uniplex /
one-layer

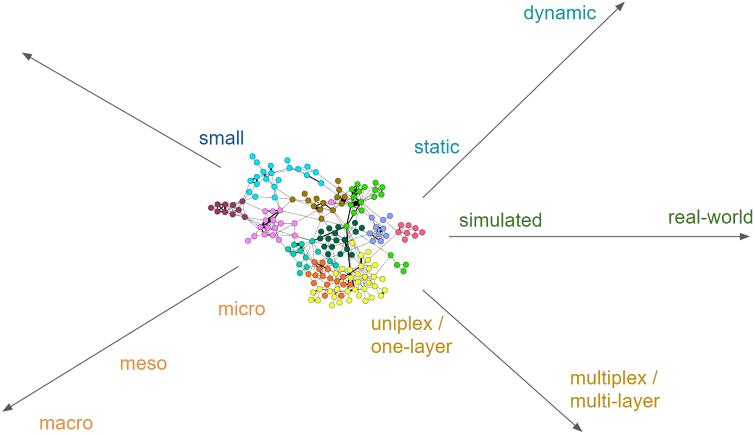
multiplex /
multi-layer

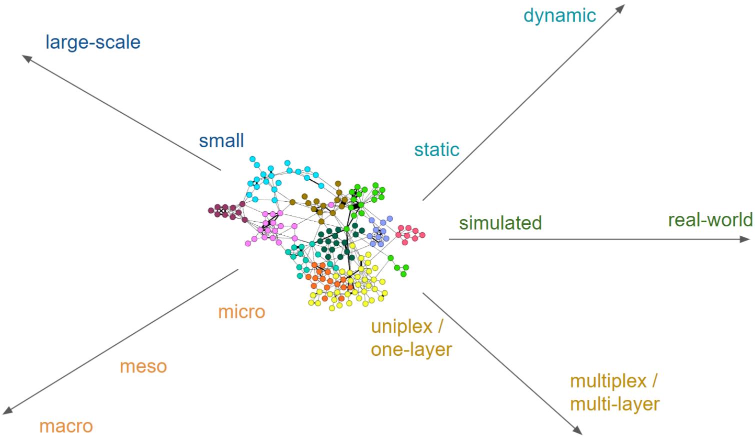

simulated
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer

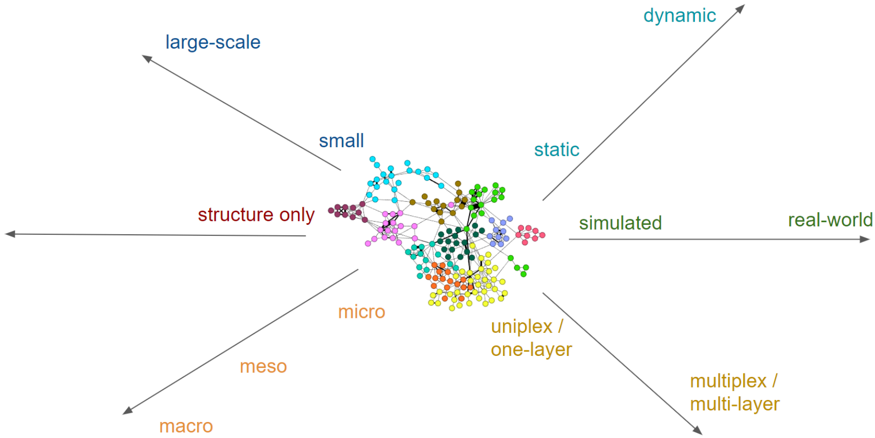

simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer

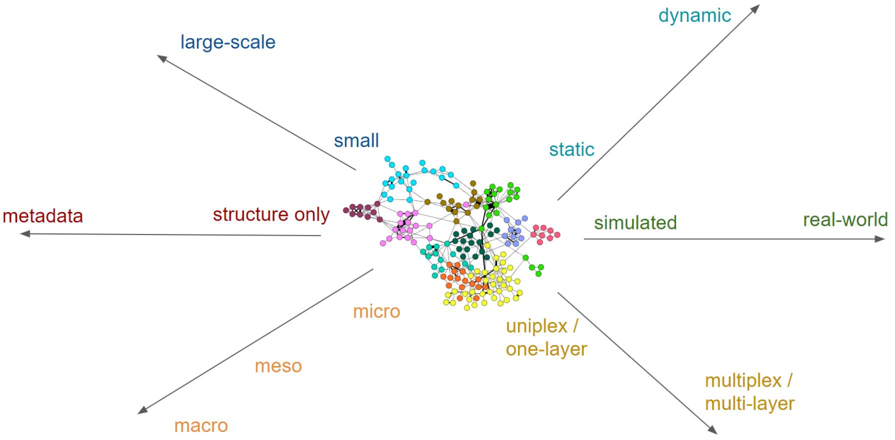

static
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer


dynamic
static
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer


dynamic
small
static
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer

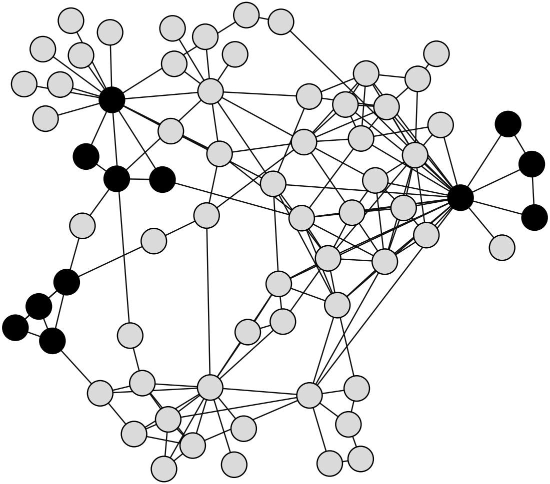

large-scale
small
dynamic
static
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer


large-scale
small
dynamic
static
structure only
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer

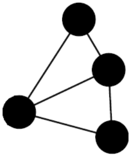

large-scale
small
dynamic
static
metadata
structure only
simulated
real-world
micro
meso
macro
uniplex /
one-layer
multiplex /
multi-layer

Network motifs

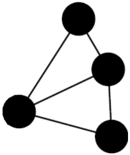
Meso scale



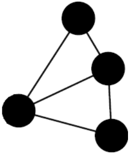
Motif

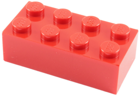


Motif



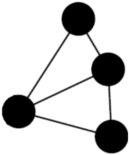
Motif





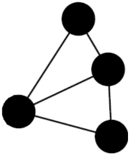
- motif: subgraph?

Motif



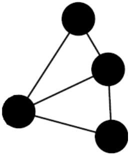
- motif: subgraph?
- motif: frequent subgraph?

Motif



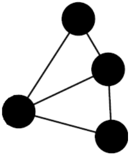
- motif: subgraph?
- motif: frequent subgraph?
- motif: surprisingly frequent subgraph?

Motif



- motif: subgraph?
- motif: frequent subgraph?
- motif: surprisingly frequent subgraph?
- motif: noteworthy subgraph?

Motif



- motif: subgraph?
- motif: frequent subgraph?
- motif: surprisingly frequent subgraph?
- motif: noteworthy subgraph?
- graphlet, graph census, . . .

Motif discovery

Step 1: Counting the subgraphs

- Input: network
- Apply subgraph detection/counting **algorithm**
- Output: frequency of each subgraph

Motif discovery

Step 1: Counting the subgraphs

- Input: network
- Apply subgraph detection/counting **algorithm**
- Output: frequency of each subgraph

Step 2: Determine motifs using one of these approaches:

1. Consider top-k most frequent subgraphs to be motifs
2. Filter a set of preselected useful subgraphs based on domain knowledge and label these as motifs
3. Compare subgraph frequencies between "similar" networks and define extreme discrepancies or similarities as motifs
4. Repeat process for a "null model" and identify motifs as the most "surprising" subgraphs

Motif discovery

Step 1: Counting the subgraphs

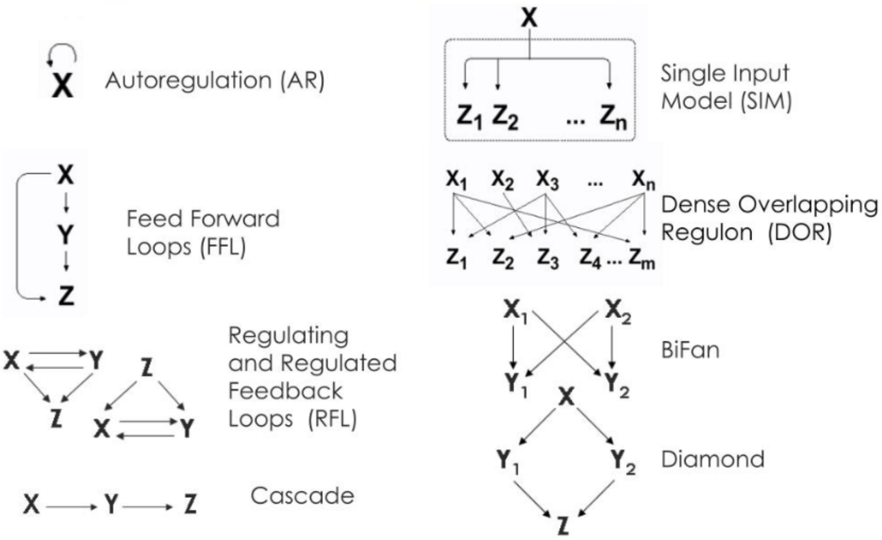
- Input: network
- Apply subgraph detection/counting **algorithm**
- Output: frequency of each subgraph

Step 2: Determine motifs using one of these approaches:

1. Consider top-k most frequent subgraphs to be motifs
2. Filter a set of preselected useful subgraphs based on domain knowledge and label these as motifs
3. Compare subgraph frequencies between "similar" networks and define extreme discrepancies or similarities as motifs
4. Repeat process for a "null model" and identify motifs as the most "surprising" subgraphs

Step 3: Reuse or interpret the results

Motifs in biological networks



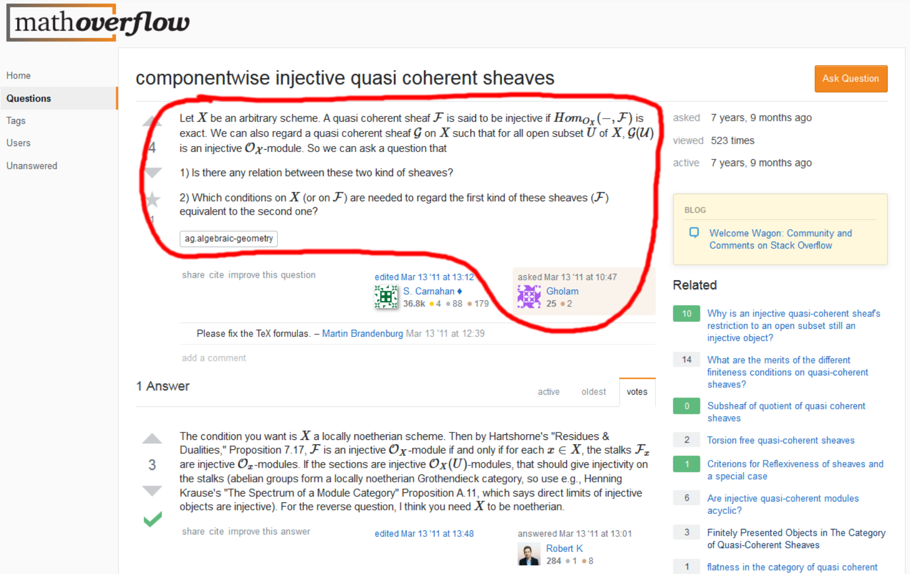
Counting motifs in multilayer temporal social networks

H.D. Boekhout, W.A. Kosters and F.W. Takes, Efficiently Counting Complex Multilayer Temporal Motifs in Large-Scale Networks, *Computational Social Networks* 6: 8, Springer, 2019.
H.D. Boekhout, W.A. Kosters and F.W. Takes, Counting Multilayer Temporal Motifs in Complex Networks, in Proceedings of the 7th International Conference on Complex Networks, *Studies in Computational Intelligence* 815: 565-577, Springer, 2018.

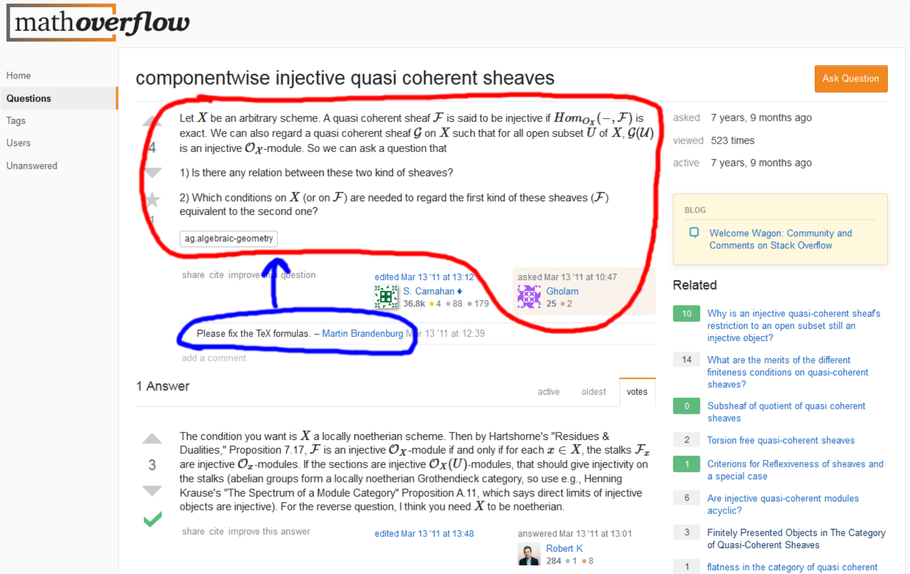
Goal

- Count motifs in **multilayer temporal** networks
- Communication from online expert knowledge exchange platform
- Nodes are users
- Three types of edges
  1. answering a question,
  2. a clarification request (commenting on a question)
  3. discussion (commenting on an answer)
- Timestamps on edges
- We define each multilayer temporal graph H as a sequence $(u_1, v_1, t_1, l_1), (u_2, v_2, t_2, l_2), \ldots (u_m, v_m, t_m, l_m)$
  Here, u_i and v_i are nodes, t_i is the timestamp of the link between these nodes and l_i is the type of link (so, the layer)

Expert exchange website



Expert exchange website



mathoverflow

Home
Questions
Tags
Users
Unanswered

componentwise injective quasi coherent sheaves

Ask Question

4

Let X be an arbitrary scheme. A quasi coherent sheaf $\mathcal{F}$ is said to be injective if $Hom_{\mathcal{O}_X}(-,\mathcal{F})$ is exact. We can also regard a quasi coherent sheaf $\mathcal{G}$ on X such that for all open subset U of X, $\mathcal{G}(\mathcal{U})$ is an injective $\mathcal{O}_X$-module. So we can ask a question that

1) Is there any relation between these two kind of sheaves?

2) Which conditions on X (or on $\mathcal{F}$) are needed to regard the first kind of these sheaves ($\mathcal{F}$) equivalent to the second one?

1

ag.algebraic-geometry

share cite improve this question

edited Mar 13 '11 at 13:12
S. Carnahan ♦
36.8k ●4 ●88 ●179

asked Mar 13 '11 at 10:47
Gholam
25 ●2

Please fix the TeX formulas. – Martin Brandenburg Mar 13 '11 at 12:39

add a comment

1 Answer

active oldest votes

3

The condition you want is X a locally noetherian scheme. Then by Hartshorne's "Residues & Dualities," Proposition 7.17, $\mathcal{F}$ is an injective $\mathcal{O}_X$-module if and only if for each $x \in X$, the stalks $\mathcal{F}_x$ are injective $\mathcal{O}_x$-modules. If the sections are injective $\mathcal{O}_X(U)$-modules, that should give injectivity on the stalks (abelian groups form a locally noetherian Grothendieck category, so use e.g., Henning Krause's "The Spectrum of a Module Category" Proposition A.11, which says direct limits of injective objects are injective). For the reverse question, I think you need X to be noetherian.

share cite improve this answer

edited Mar 13 '11 at 13:48

answered Mar 13 '11 at 13:01
Robert K
284 ●1 ●8

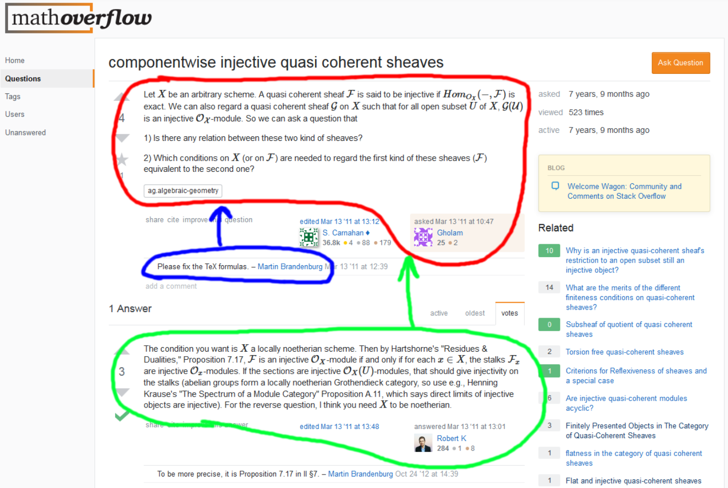
asked 7 years, 9 months ago
viewed 523 times
active 7 years, 9 months ago

BLOG
Welcome Wagon: Community and Comments on Stack Overflow

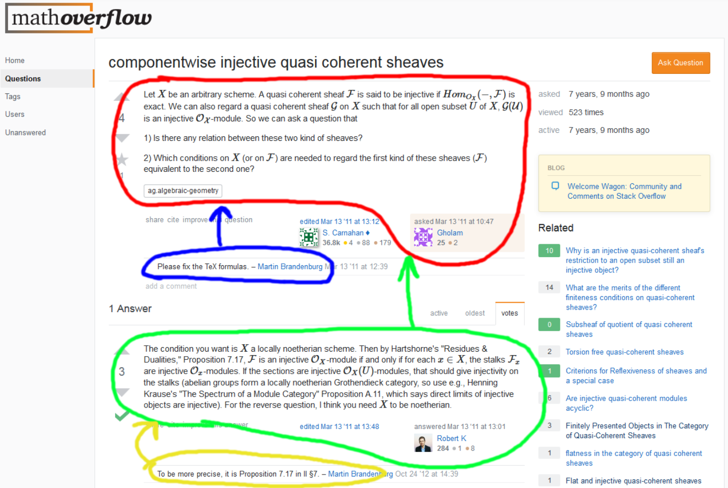
Related

10 Why is an injective quasi-coherent sheaf's restriction to an open subset still an injective object?
14 What are the merits of the different finiteness conditions on quasi-coherent sheaves?
0 Subsheaf of quotient of quasi coherent sheaves
2 Torsion free quasi-coherent sheaves
1 Criterions for Reflexiveness of sheaves and a special case
6 Are injective quasi-coherent modules acyclic?
3 Finitely Presented Objects in The Category of Quasi-Coherent Sheaves
1 flatness in the category of quasi coherent

Expert exchange website



Expert exchange website



Multilayer temporal motifs

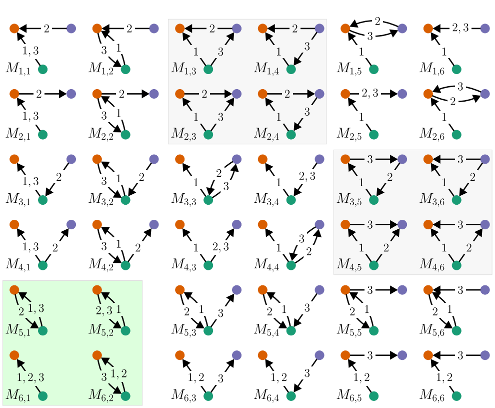
We define r-nodes, s-edges, δ-temporal, λ-layered motifs as

- a sequence of s edges,
  $M = ((u_1, v_1, t_1, l_1), (u_2, v_2, t_2, l_2), \ldots, (u_s, v_s, t_s, l_s))$, with $u_i, v_i \in E$;
- of δ duration, i.e., $t_1 < t_2 < \ldots < t_s$ and $t_s - t_1 \leq \delta$;
- ranging over at most λ different layers;
- and having r nodes.

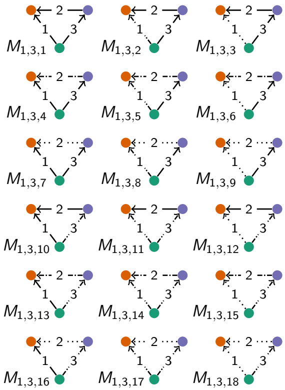
Problem statement

Given set values for r, s, δ and λ and a multilayer temporal graph H, compute the number of occurrences of each motif.

Motif types (size 3)

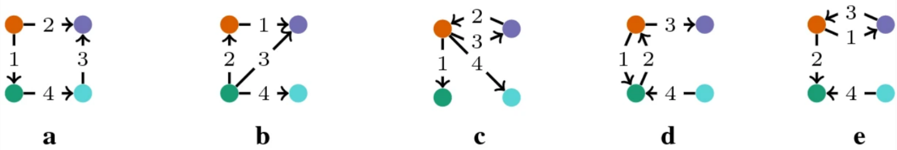


(a) The 36 2,3-node, 3-edge δ-temporal motifs (figure from Paranjape et al., 2017).



(b) 18 of the 27 3-layer variants of $M_{1,3}$.

Motif types (size 4)



Types of 4-node, 4-edge, temporal motifs. **a** Square, **b** Tailed-triangle, **c** Star, **d** Mid-Path, **e** Head-Path

Figure: The 624 Square (48), Tailed-Triangle (192), Star (96), Mid-Path (96) and Head-Path (192) motifs; 624 in total.

Temporal motif counting

Paranjape et al. (2017) introduced 3 counting algorithms

1. General, based on underlying static motifs (2-node motifs);
2. Star, based on 'center' node;
3. Triangle, based on edges involved in the most triangles.

Time complexity in the order of the size of the input, i.e., $O(m)$

Temporal motif counting

Paranjape et al. (2017) introduced 3 counting algorithms

1. General, based on underlying static motifs (2-node motifs);
2. Star, based on ‘center’ node;
3. Triangle, based on edges involved in the most triangles.

Time complexity in the order of the size of the input, i.e., $O(m)$

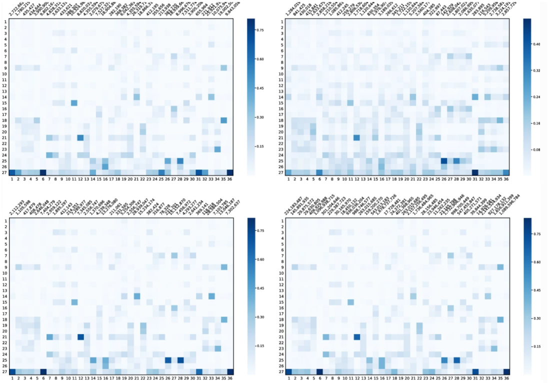
We extend these to deal with *multilayer* and *partially timed* motifs.
Implemented as an extension of SNAP

Datasets

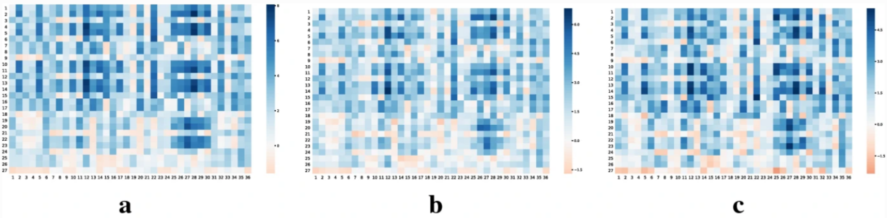
Table: Network dataset statistics.

Dataset	**Nodes**	**Edges**	**Edges**$_{static}$	λ	**deg**$_{max}$
EMAIL-EU-CORE	985	24,929	24,929	2	345
MATH-OVERFLOW	24,759	390,441	228,215	3	2,172
FACEBOOK/WOSN	63,792	2,401,228	1,592,562	2	1,100
ASK-UBUNTU	157,222	726,661	544,774	3	5,401
SUPER-USER	192,409	1,108,739	854,377	3	14,294
STACK-OVERFLOW	2,584,164	47,903,266	34,901,115	3	44,065

Results



Results MathOverflow



Differences between the motif footprints of the three most distinct expert knowledge exchanges. **a**–**c** each denotes a pair (MATH-OVERFLOW vs. DATASETX). For a multilayer motif (cell), color is proportional to the difference, where blue denotes that the motif is more dominant in MATH-OVERFLOW, and analogously orange in DATASETX. Color gradient is proportional to the $\log_2$ difference. Values between parentheses denote the average difference between all 27×36 column-normalized counts. **a** MATH-OVERFLOWvs. STACK-OVERFLOW (0.50), **b** MATH-OVERFLOW vs. SUPER-USER (0.43), (c) MATH-OVERFLOW vs. ASK-UBUNTU (0.39)

Findings

- Non-appearing motifs show the non-building-blocks :-)
- Reciprocated question-answer links (layer 1) are rare; clear difference between experts and novice users.
- Reciprocation does occur in the discussion layer (layer 3), also in triangles.
- The abovementioned effect is even stronger in StackOverflow (more of a helpdesk than an expert knowledge exchange)
- The computer science communities (StackOverflow, Ask-Ubuntu, Super-User) have most communication in the comment-on-answer layer. In Math-overflow there is much more question-answer and question-commenting activity (what does this say about the difference between CS and math?)

Enumerating motifs in multiplex corporate networks

F.W. Takes, W.A. Kosters, B. Witte and E.M. Heemskerk, Multiplex Network Motifs as Building Blocks of Corporate Networks, *Applied Network Science* 3: 39, Springer, 2018.
F.W. Takes, W.A. Kosters and B. Witte, Detecting motifs in Multiplex Corporate Networks, in Proceedings of the 6th International Conference on Complex Networks, *Studies in Computational Intelligence* 642: 502-515, Springer, 2017.

Corporate networks

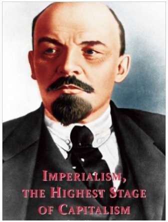
- **Nodes** are organizations/firms/companies/corporations

Corporate networks

- **Nodes** are organizations/firms/companies/corporations
- **Links** represent, e.g., trade, loans, ownership, interlock
- **Ownership**: firm A owns (part of) the shares of firm B and can thus control it
- **Board interlock**: there is a relationship between firms because they share a board member or director

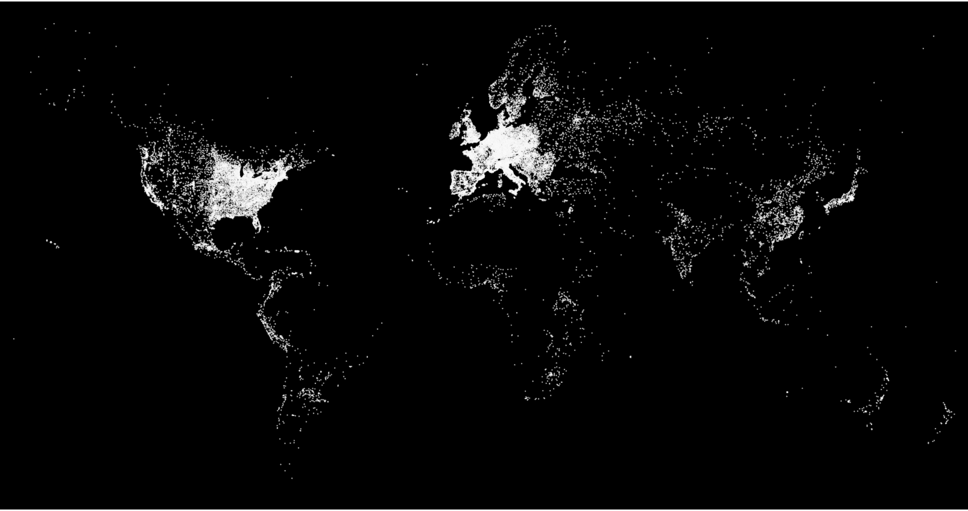
Corporate networks

- **Nodes** are organizations/firms/companies/corporations
- **Links** represent, e.g., trade, loans, ownership, interlock
- **Ownership**: firm A owns (part of) the shares of firm B and can thus control it
- **Board interlock**: there is a relationship between firms because they share a board member or director



Vladimir I. Lenin, *Imperialism, The Highest Stage of Capitalism*, 1916
"... a personal union, so to speak, is established between the banks and the biggest industrial and commercial enterprises, the merging of one with another through the **acquisition of shares**, *through the* **appointment of** *bank* **directors** *to the Supervisory Boards (or Boards of Directors) of industrial and commercial enterprises, and vice versa."*

Corporations



CORPNET

F.W. Takes and E.M. Heemskerk, Centrality in the Global Network of Corporate Control, *Social Network Analysis and Mining* 6(1): 1-18, 2016.

Multiplex corporate networks

- One set of nodes (firms)

Multiplex corporate networks

- One set of nodes (firms)
- Multiple sets of edges
  - Ownership (directed)

    
  - Board interlocks (undirected)

Multiplex corporate networks

- One set of nodes (firms)
- Multiple sets of edges
  - Ownership (directed)

  - Board interlocks (undirected)

- Interlayer assortativity: different types of edges are related to each other.

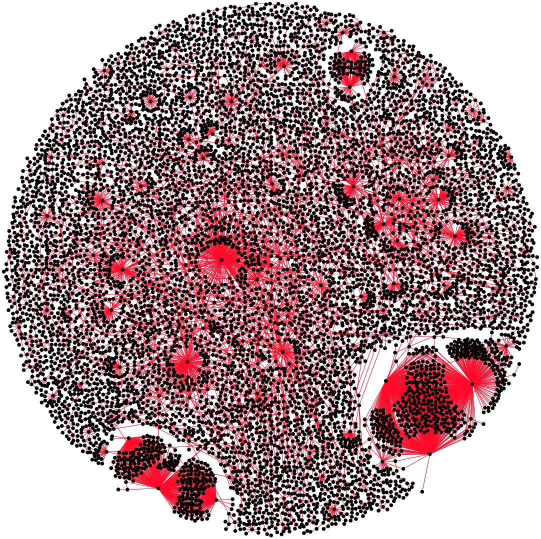
Multiplex corporate networks

- One set of nodes (firms)
- Multiple sets of edges
  - Ownership (directed)

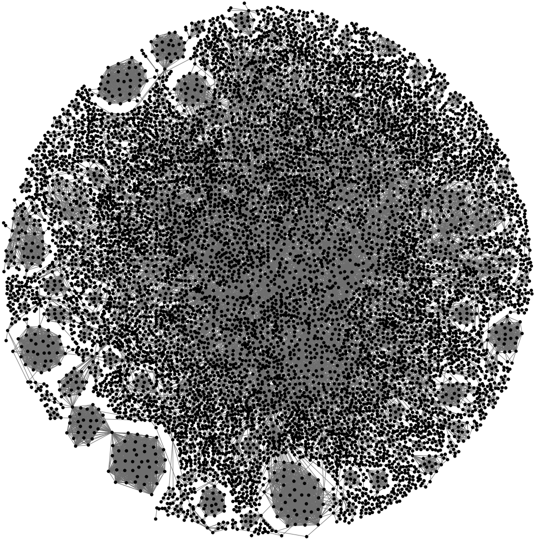
    
  - Board interlocks (undirected)

    
- Interlayer assortativity: different types of edges are related to each other. Frequently (5.9% of links): 
- Challenge: enumerate **multiplex motifs**
- Network size: 75 224 nodes, 195 073 edges

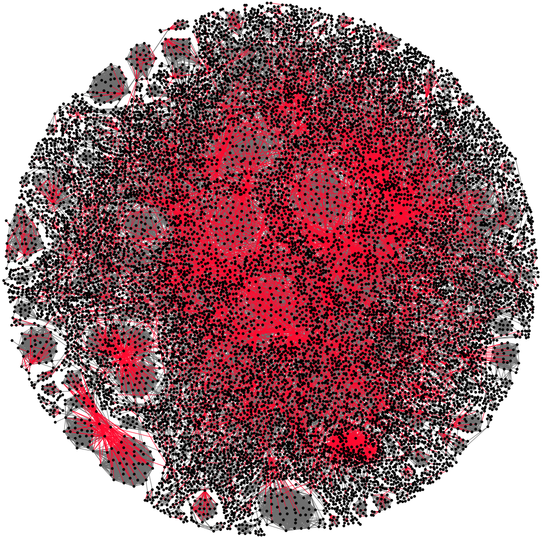
Ownership network



Board interlock network



Multiplex network



Data

Table: Division of firms over economic sectors

Sector	Ownership		Board interlock		Multiplex	
Bank	474	1.25%	865	1.41%	972	1.29%
Financial	4 648	12.32%	6 250	10.21%	8 338	11.08%
Foundation/Research	55	0.14%	51	0.08%	88	0.12%
Industrial	32 350	85.75%	53 767	87.84%	65 484	87.05%
Insurance	19	0.05%	26	0.04%	34	0.05%
Mutual/Pension Fund	112	0.30%	175	0.29%	213	0.28%
Private Equity	29	0.08%	30	0.05%	37	0.05%
Public Authority	22	0.06%	31	0.05%	41	0.05%
Venture Capital	15	0.04%	14	0.02%	17	0.02%

Null model

- Stub matching model
- Challenge: deal with interlayer assortativity (overlap between layers)
- Modeling each *layer* will not realize 5.9% overlap between layers

Null model

- Stub matching model
- Challenge: deal with interlayer assortativity (overlap between layers)
- Modeling each *layer* will not realize 5.9% overlap between layers
- Instead, we model each *combination of layers* separately
  For L layers, we have $2^L - 1$ models
  0. No link
  1. Ownership link
  2. Board interlock link
  3. Ownership and board interlock link

Null model

- Stub matching model
- Challenge: deal with interlayer assortativity (overlap between layers)
- Modeling each *layer* will not realize 5.9% overlap between layers
- Instead, we model each *combination of layers* separately
  For L layers, we have $2^L - 1$ models
  0. No link
  1. Ownership link
  2. Board interlock link
  3. Ownership and board interlock link
- Combine modeled layers into multiplex network
- Sample a large number of models, and count its subgraphs

Motif significance testing

1. Subgraph **ratio**:

$$\frac{\text{frequency in data}}{\text{frequency in random model samples}}$$

2. Subgraph **concentration**:

$$\frac{\text{frequency in data}}{\text{frequency of all patterns of that size}}$$

- Cut-off value for ratio (5) and concentration (0.01) determines which subgraphs (patterns) of size $k = 3$, $k = 4$ and $k = 5$ are **motifs**
- Implemented as an extension of SubEnum

Results

Table: Patterns and motifs per network

	Pattern size				Motif size			
	3	**4**	**5**	**All**	**3**	**4**	**5**	**All**
Ownership	11	63	391	465	3	4	6	13
Board interlock	2	6	21	29	0	2	10	12
Multiplex	58	1 132	21 858	23 048	14	48	73	135

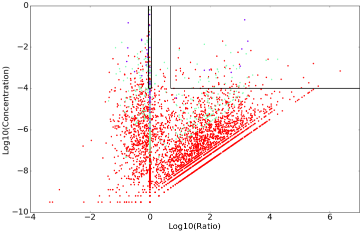


Figure: Ratio (horizontal axis) vs concentration (vertical axis) for all patterns. Top right box indicates cut-off values. Patterns of size 3 in blue, size 4 in green and size 5 in red.

Highlighted motif

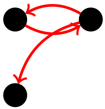


Figure: Ownership motif of size 3 with ratio 26 and concentration 0.026. Crossholdings (common in Germany). Dominated by industry sector.

Highlighted motifs

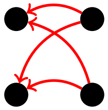


Figure: Multiplex motif of size 4 with ratio 2024 and concentration 0.351. Two joint ventures. Dominated by financial sector (56%).

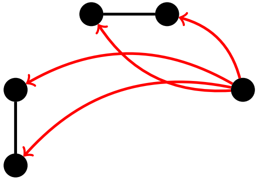


Figure: Multiplex motif of size 5 with ratio 113 400. Two investments into two firms governed by the same director. Dominated by "Mutual & Pension Fund" sector (14%).

Results

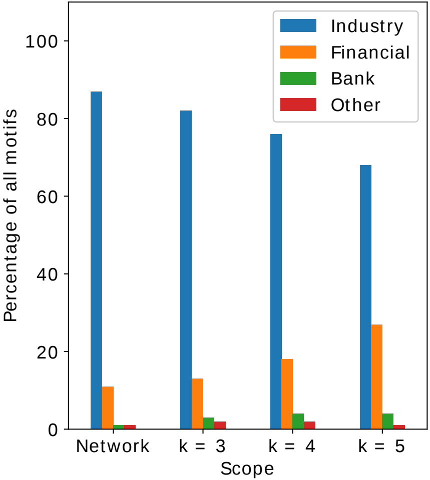


Figure: Division of firms over sectors for full network and different motifs.

Building blocks . . .

Conclusion

- Meso level patterns: communities and **motifs**
- Trade-off between enumerating and counting motifs
- Counting can also be done in a "complex" network (timed, layered, attributed)
- Motifs can help characterize complex communication patterns in online communities
- Certain motifs in corporate networks appear specifically in certain industry sectors
- Motifs in corporate networks reveal the role of the financial sector

Thank you!

- Questions?
- https://franktakes.nl
- https://computationalnetworkscience.org
- https://corpnet.uva.nl

NetSci Dutch Network Science Society

Leiden Complex Networks Network LCN2

Figure: http://www.netsci.nl http://lcn2.leidenuniv.nl