DATA ANALYSIS II

Multilayer Social Networks III
Visualization and Community Detection
2021/22

Outline

- Visualization - examples
- Community detection

Visualization – examples based on AUCS data

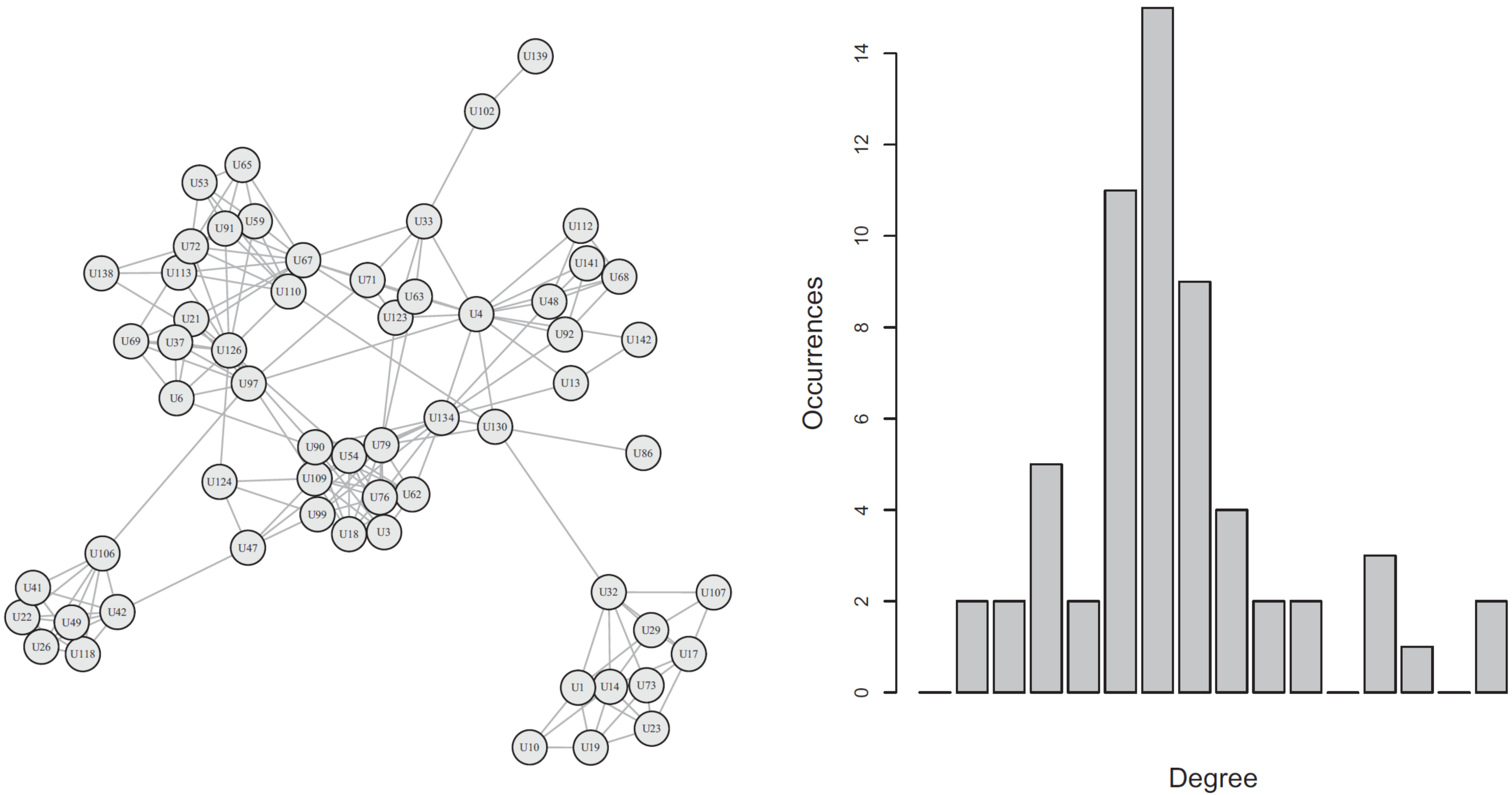
- AUCS data were collected at a university research department and include five online and offline layers. The population of the study consists of 61 employees (out of the total number of 142) who decided to join the survey, including professors, postdoctoral researchers, PhD students, and administrative staff. The role and anonymized research group of each actor is also specified as an attribute.

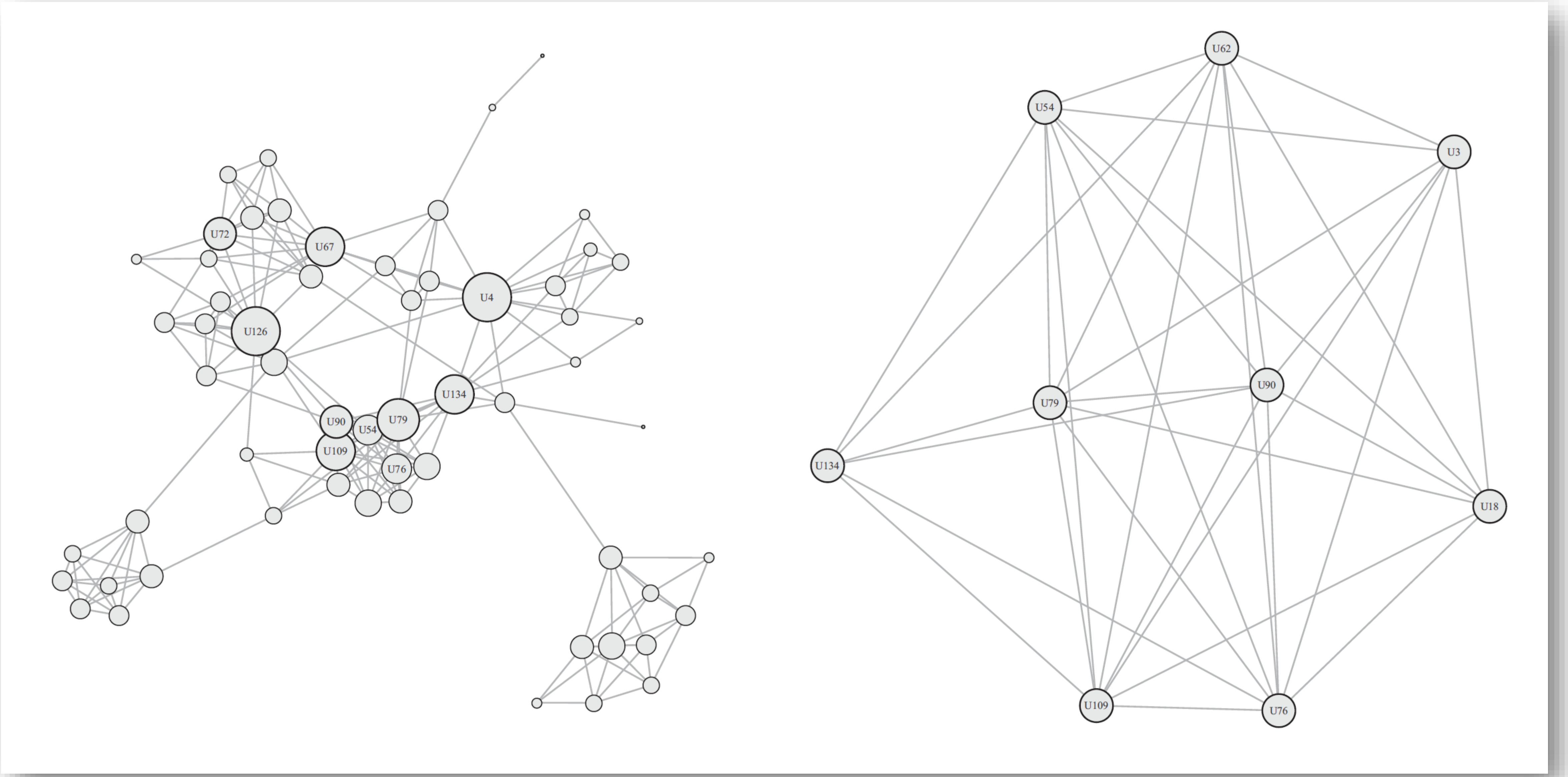
- It consists of 5 layers (facebook, lunch, coauthor, leisure, work), 61 nodes, 620 edges (there is a need to ask for the dataset).

https://drive.google.com/drive/folders/0B_M5Zh3gg4LkNWZ5WmpoRlJhMVE

Single Network Visualization

- Four typical visualizations of a single network:
  - focus on structure
  - focus on metrics
  - augmented visualization
  - simplified visualization (e.g., 6-core of the network).

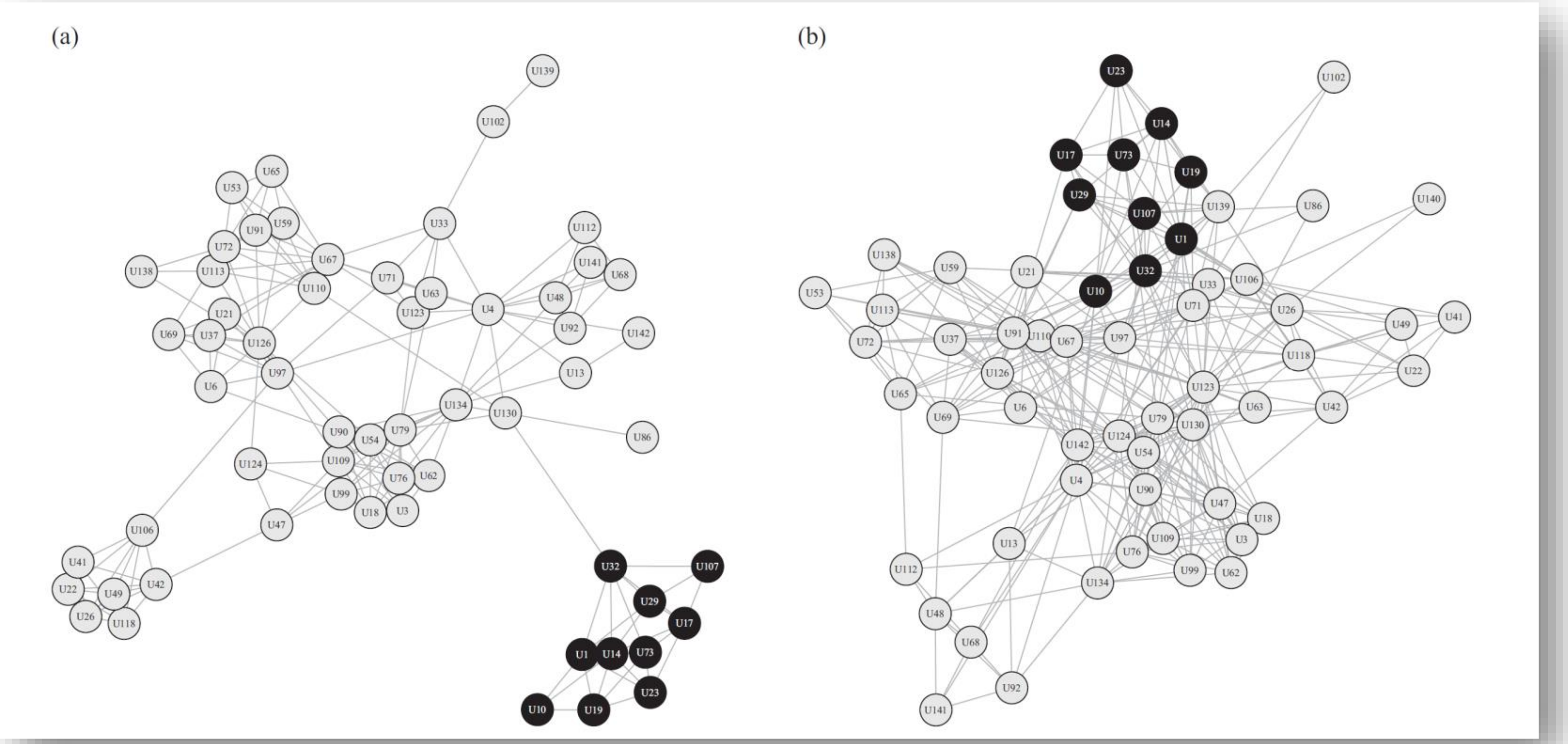

U139
U102
U65
U53
U59
U91
U33
U112
U72
U67
U141
U68
U138
U113
U71
U110
U63
U48
U4
U123
U21
U92
U142
U69
U37
U126
U97
U13
U6
U134
U130
U86
U90
U79
U54
U109
U124
U62
U76
U99
U3
U18
U47
U106
U41
U32
U107
U42
U22
U49
U29
U26
U118
U17
U1
U14
U73
U23
U10
U19
Occurrences
0
2
4
6
8
10
12
14
Degree



U72
U67
U4
U126
U90
U54
U79
U134
U109
U76
U62
U54
U3
U79
U90
U134
U18
U109
U76

One layer x flattening

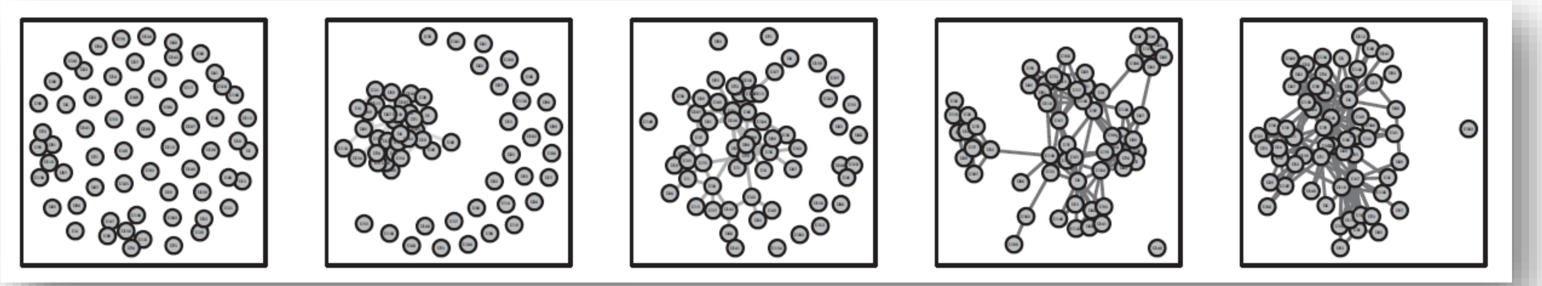
- One/five layers of complexity
- a single layer, *lunch*
- the full flattened AUCS network

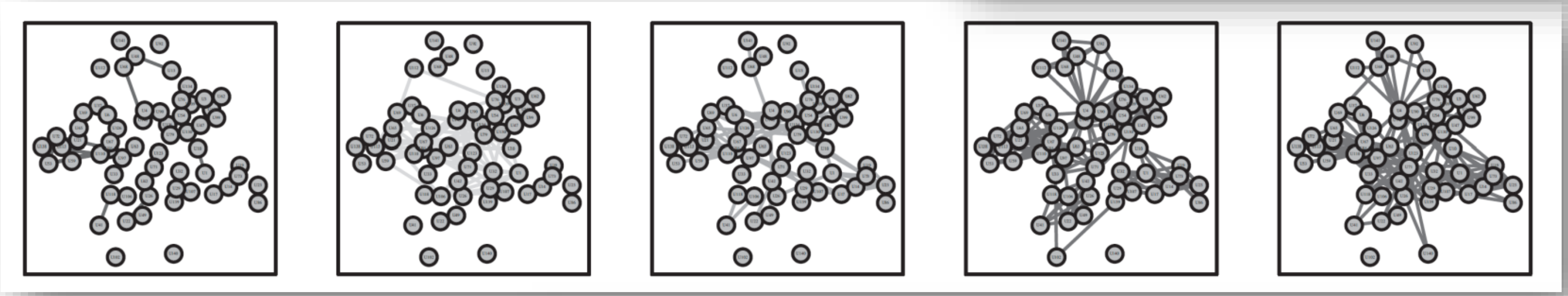


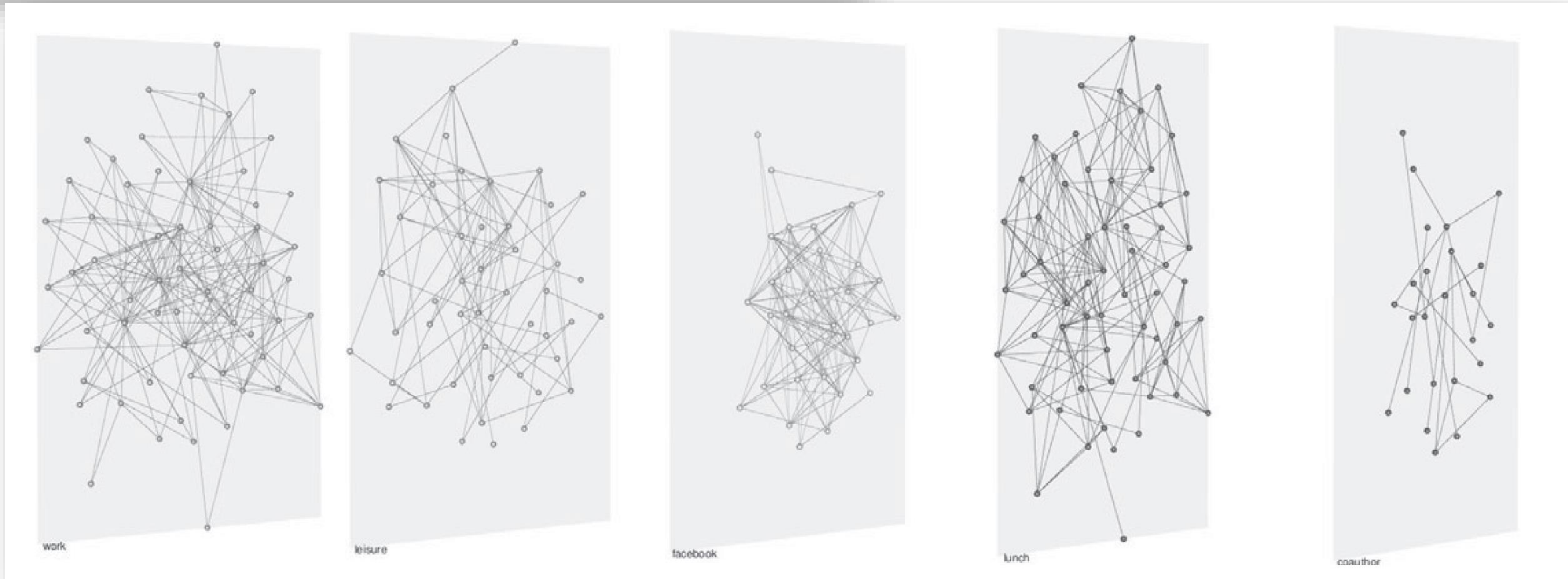
Structure density

- Comparing the two visualizations, we can see how the clear structure of the lunch network becomes more blurred and confused if we take connections from all the layers into consideration.
    - As an example, a clearly visible cluster on the left-hand side layer is highlighted (black nodes).
    - The same nodes are also black-marked in the flattened graph on the right, and we can see that the cluster has been partially attracted toward the center of the figure and that some of its nodes are now more connected to other nodes outside the cluster.

Layer slicing







The same layout for each layer

- Two alternative visualizations can be used. Both methods slice the network into its composing layers. To simplify a comparison between the layers, the nodes have been placed using the same layout in each slice.

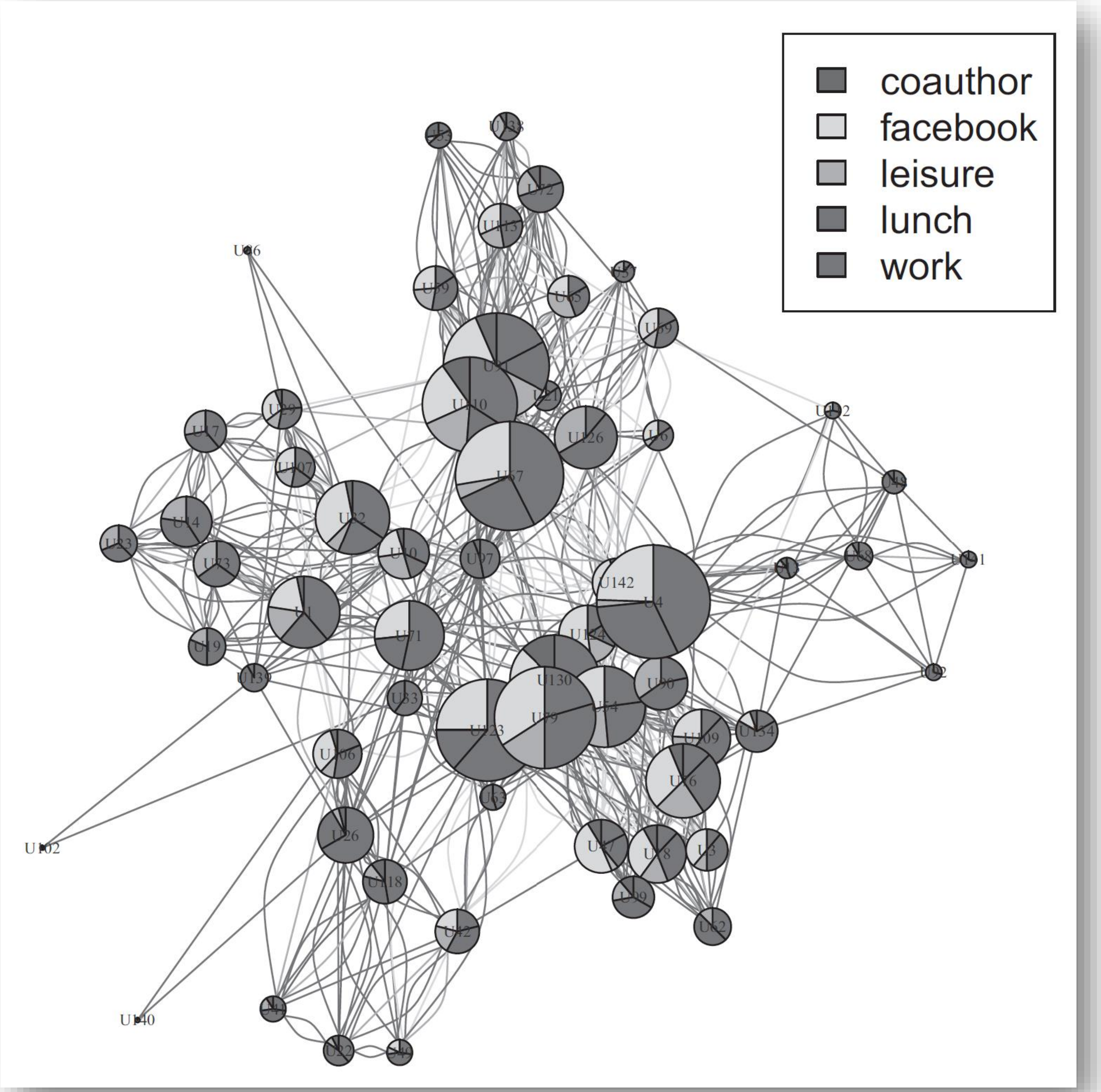
- A very similar approach can be used, as it is done by the software *MuxViz* to obtain an interactive 3-dimensional visualization of a multilayer network. It is also called a 2.5-dimensional representation, because it is made of 2-dimensional planes.

- MuxViz: https://muxviz.wordpress.com/info/

Augmented network

AUCS network augmented with actor degrees, visualized using *igraph*

https://igraph.org



Structure + measures

- Multilayer metrics can be used to increase the information contained within the graphical visualizations
- Every actor contains information about its degree (size of the circle representing the actor) and its neighborhood in the various layers (pie chart).
- It combines the layout structure defined on the flattened network with metrics computed both on the flattened network (degree) and on the multilayer network (neighborhood).

Other R and Python libraries

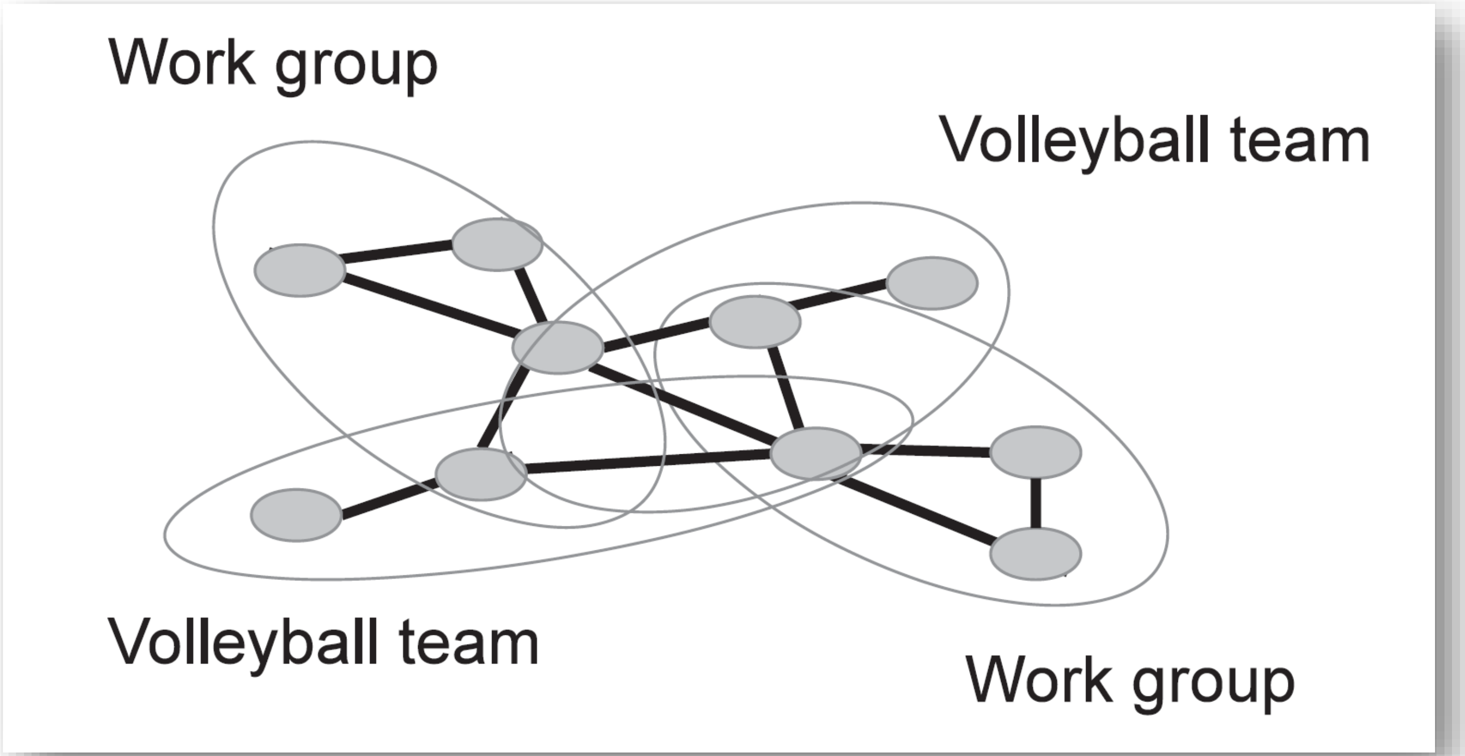
- http://multilayer.it.uu.se/software.html
- Py3plex toolkit for visualization and analysis of multilayer networks
  - https://link.springer.com/article/10.1007/s41109-019-0203-7

Community detection

- Freeman defines groups (communities) as relatively small, informal, and formed by close personal ties. These considerations lead one to expect greater homogeneity among members of the same group.

- Four general properties of *cohesive subgroups* have influenced most of the social network formalizations:
  - the mutuality of ties;
  - the closeness or reachability of subgroup members;
  - the frequency of ties among members;
  - the relative frequency of ties among subgroup members compared to nonmembers.

Freeman, Linton C. 1996. Some antecedents of social network analysis. *Connections*, 19(1), 39–42.



- In large networks, a dense core surrounded by smaller extensions can be the result of several overlapping clusters.
- Peripheral nodes are the ones belonging to one or few clusters, whereas central ones belong to many.

Three main approaches

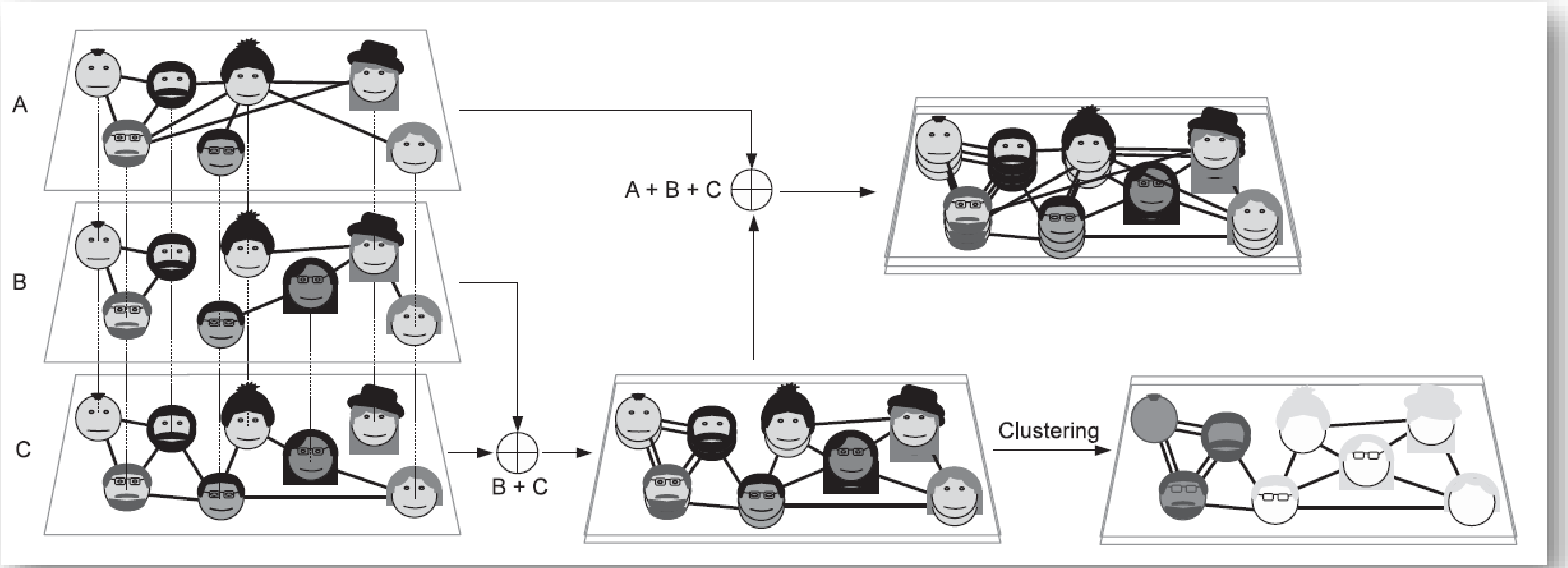
- Applying existing algorithms to simplified (e.g., flattened) data that have been first reduced to a single-layer network, or at least to a less entangled network.
- An approach considering one layer at a time, and after single-layer communities have been identified, they are merged into larger structures spanning multiple layers.
- Extensions of the most common single-layer network methods to a multilayer model, including modularity optimization, random walks, spectral clustering, and label propagation.
- *Despite the availability of multilayer approaches, there might be practical reasons for researchers to use flattened network methods.*

Simplification – Flattening Based on Relevance

- Away to simplify a multilayer network is to use one of the flattening approaches. Then, any community detection algorithm can be used (e.g., Louvain).

- *Problem: The analysis of the combination of all layers, that is, the flattened network, does not have to reveal any interesting patterns (communities), as there can be too many edges.*

- A combination of some layers can be chosen manually whenever we have some specific qualitative knowledge of the network indicating which combination might be relevant (not always available).



Multiple layers are combined to obtain a single network, then traditional community detection algorithms can be used. However, communities may appear when a specific subset of the layers is used (B + C) and disappear or become less dense when less or more layers are used (like A, or A + B + C).

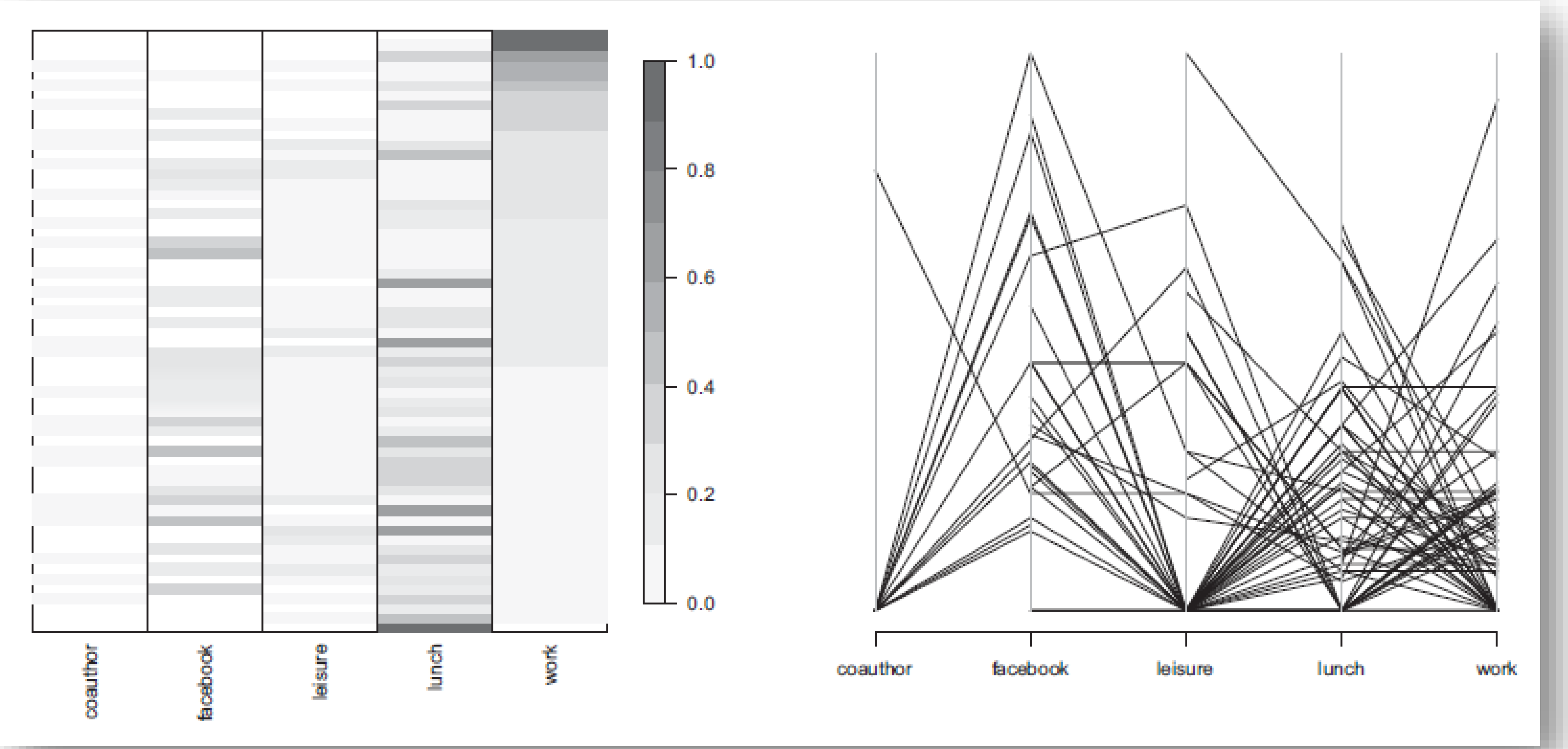
Local Simplification Process

- Given a measure $r(a, l)$ indicating how much a layer l is relevant for actor a, and a threshold θ, for each layer l and for each pair of actors a_1, a_2, we keep an edge between them on this layer if and only if
  - an edge exists,
  - $r(a_1, l) \geq \theta$,
  - $r(a_2, l) \geq \theta$.

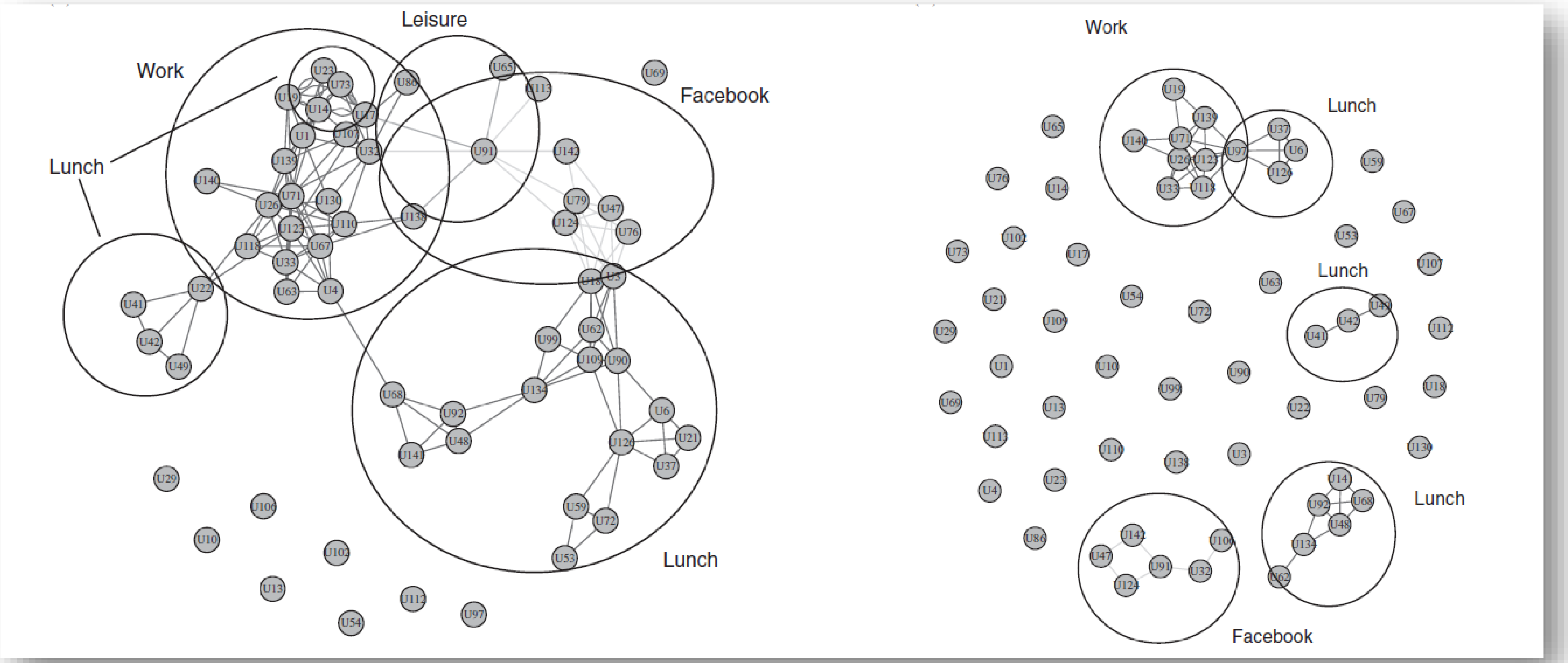
This indicates that the dyadic relation between a_1 and a_2 is considered an important one by both actors and should thus be preserved.

Layer Relevance

- Different measures emphasizes the relationships between different layers.
  - The value of relevance for each actor on each layer of the AUCS data set.
  - Number of neighbors for each actor on each layer.

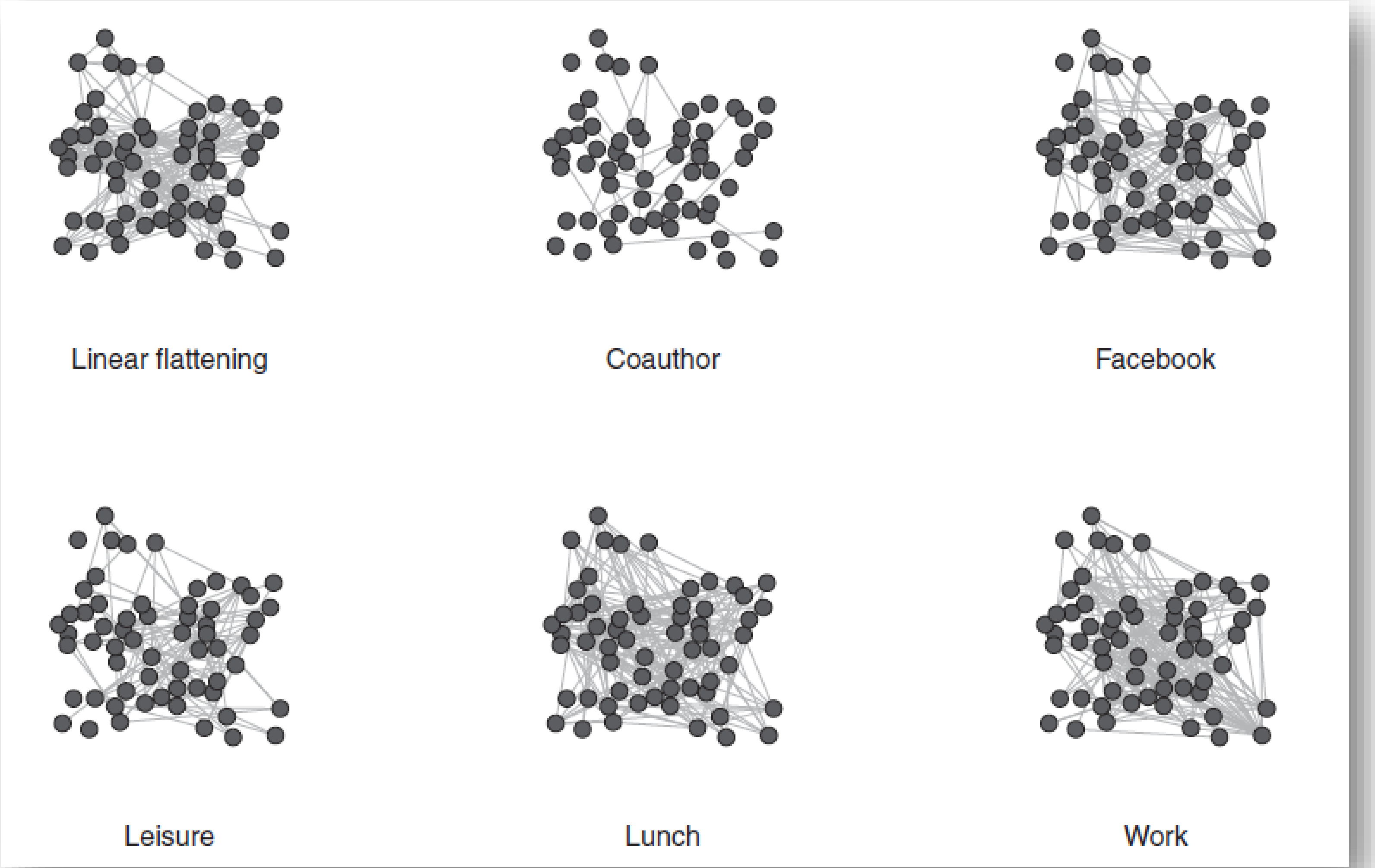


Local Simplification

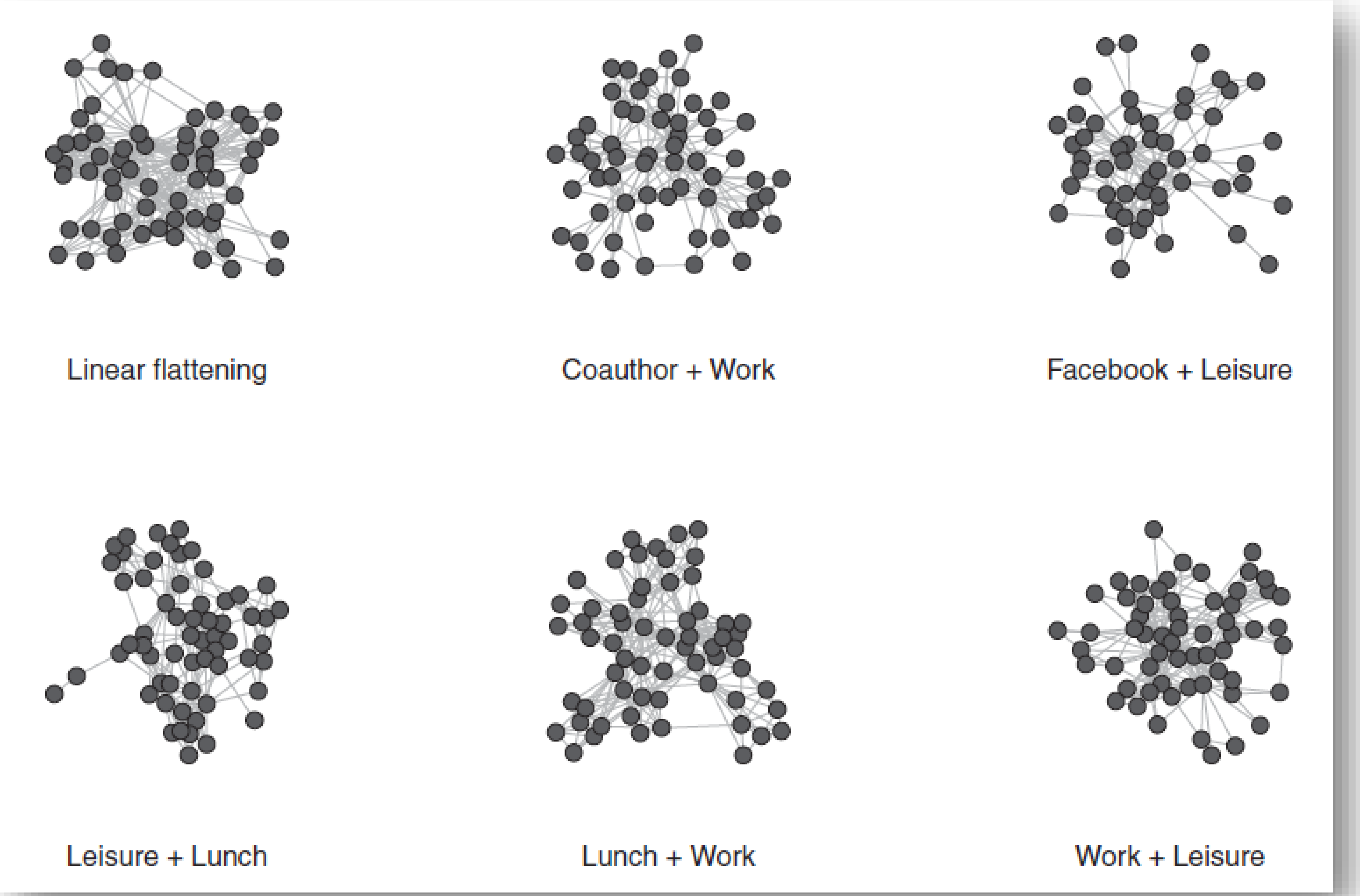


- *Relevance* ≥ *0.6* and *exclusive relevance* ≥ *0.3*.

AUCS: Particular layers



AUCS: Various layer combinations



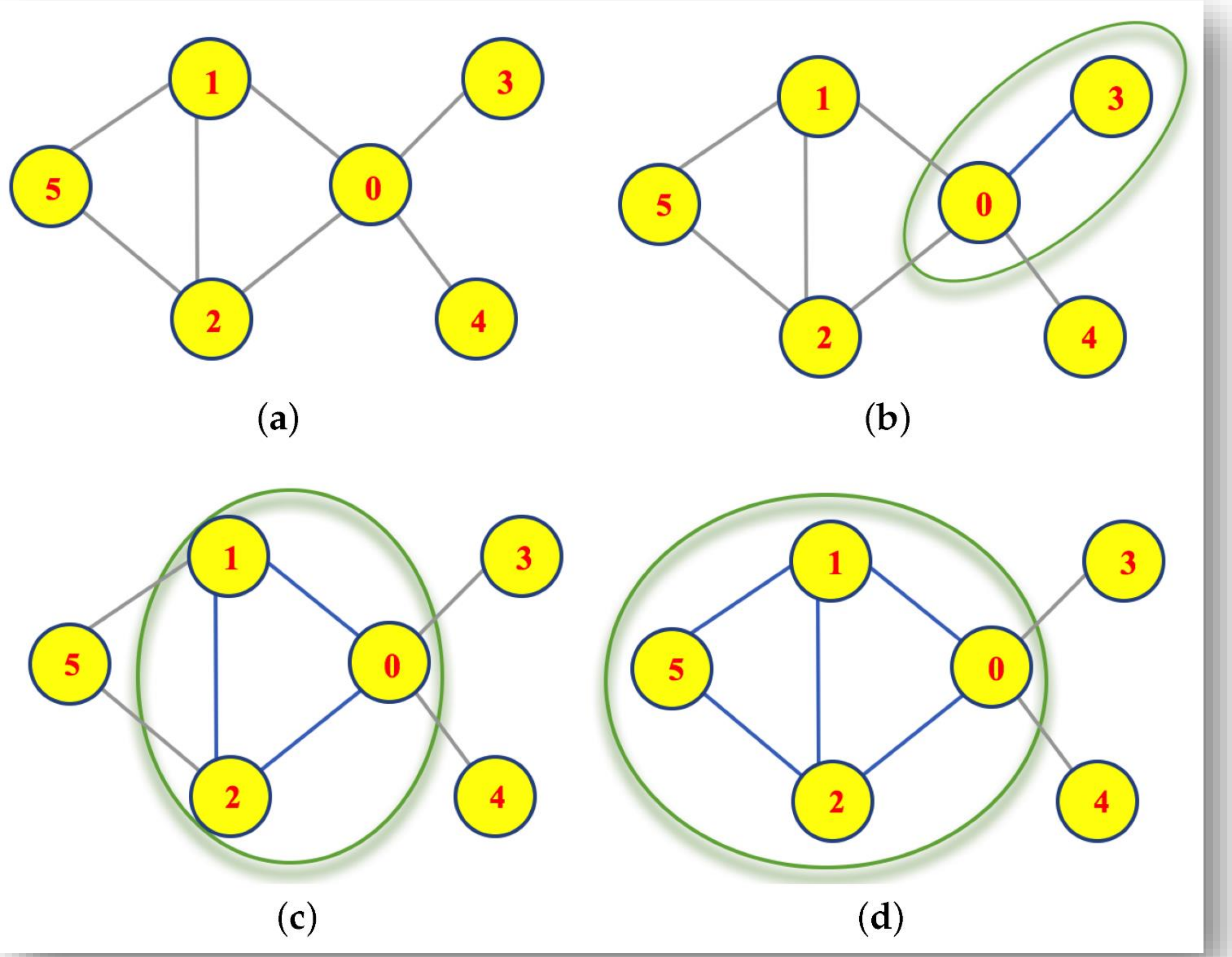
Combination of Single-layer Communities

- These approaches are based on the idea of detecting communities on single layers, using existing methods, and then aggregating them into larger structures spanning multiple layers.

- A simple approach can be based on cliques.
  - **Definition (Clique)** A clique is a set of nodes connected to all other nodes in the clique.
  - **Definition (Maximal clique)** A maximal clique is a clique that is not contained in a larger clique.
  - **Definition (Quasi-clique)** A quasi-clique is a set of nodes where each node is connected to at least a fraction γ of the other nodes in the quasi-clique.

Quasi-cliques

- A simple graph with 0.5-quasi-cliques of different sizes containing the node 0

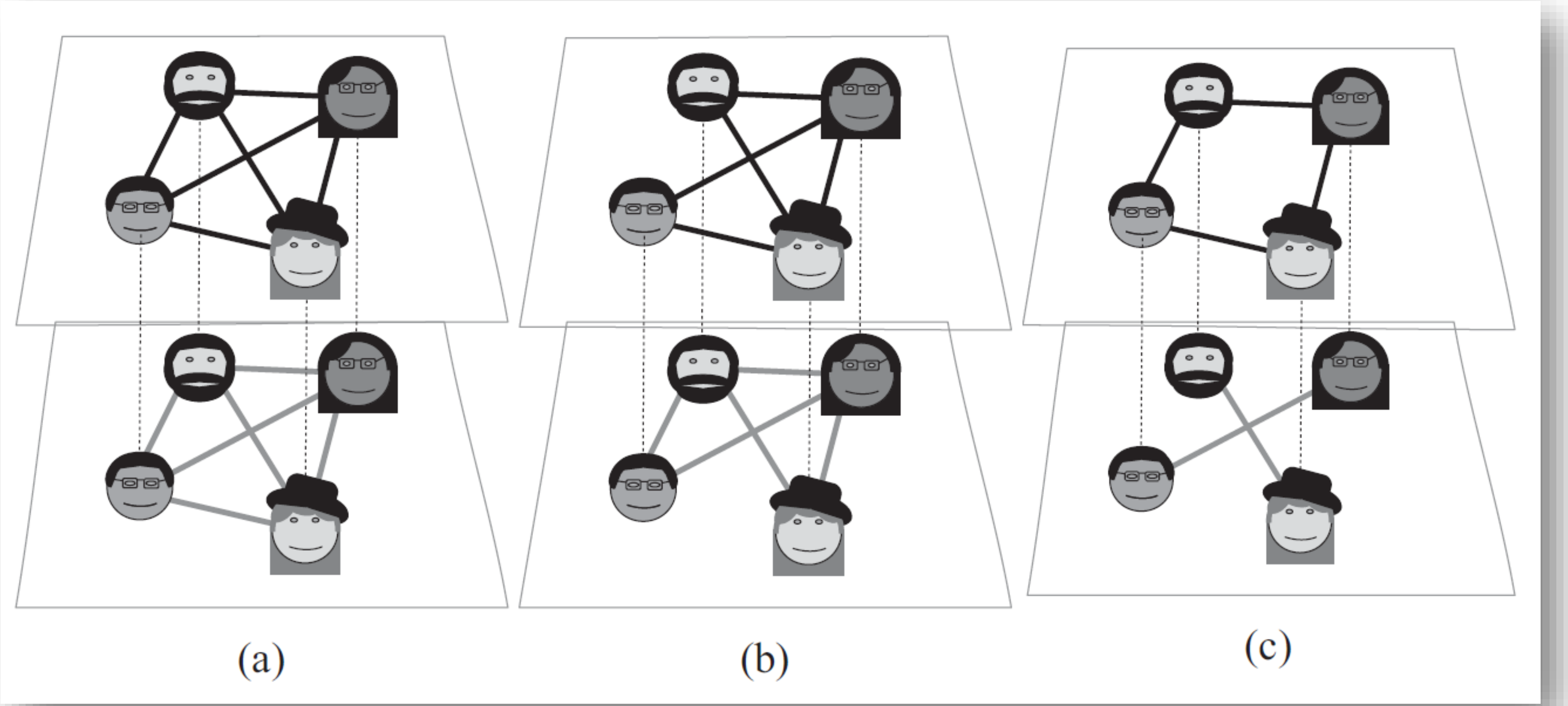


(a) (b) (c) (d)

Cliques for multiplayer networks

- **Definition (Multilayer clique)** Given a set of layers L, a multilayer clique is a set of actors connected to all other actors in the clique on each of these layers.
- **Definition (Multilayer quasi-clique)** Given a set of layers L; a multiplayer quasi-clique is a set of actors where each actor is connected to at least a fraction γ of the other actors in the quasi-clique on at least a fraction λ of the layers composing the multilayer network.
- When $\lambda = 1 / |L|$, cliques may emerge from the contribution of different layers where no clique structure would be otherwise observable.

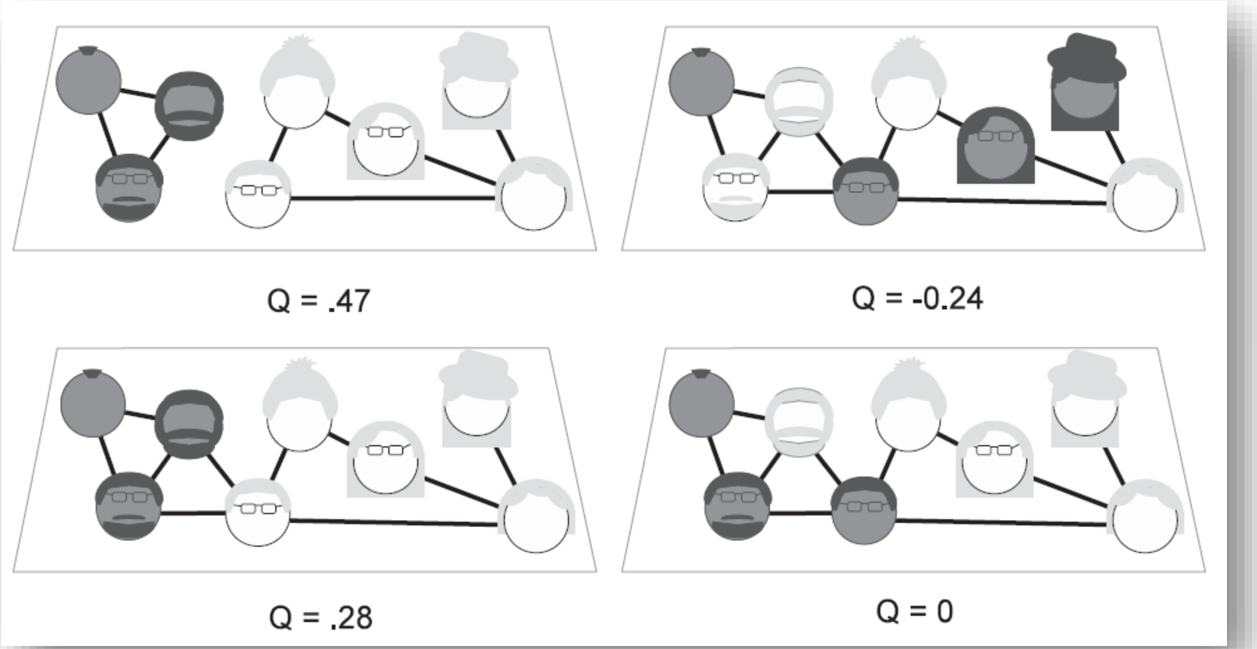
Example



Multilayer cliques: (a) a multilayer clique, (b) a quasi-clique with $\lambda = 1$ and $\gamma = .5$, and (c) a quasi-clique with $\lambda = 1/2$ and $\gamma = 1$.

Multilayer Modularity Optimization

- Modularity Q is a measure of how well actors can be separated into dense and independent components. Informally, modularity is a measure of how well subgroups are separated with the network.



$$Q = \frac{1}{2m} \sum_{ij} \left(a_{ij} - \frac{k_i k_j}{2m} \right) \delta(\gamma_i, \gamma_j)$$

Extension of modularity for multilayer networks

- If we assign each node to one single group, thanks to the multiple layers, actors can belong to different groups at the same time.

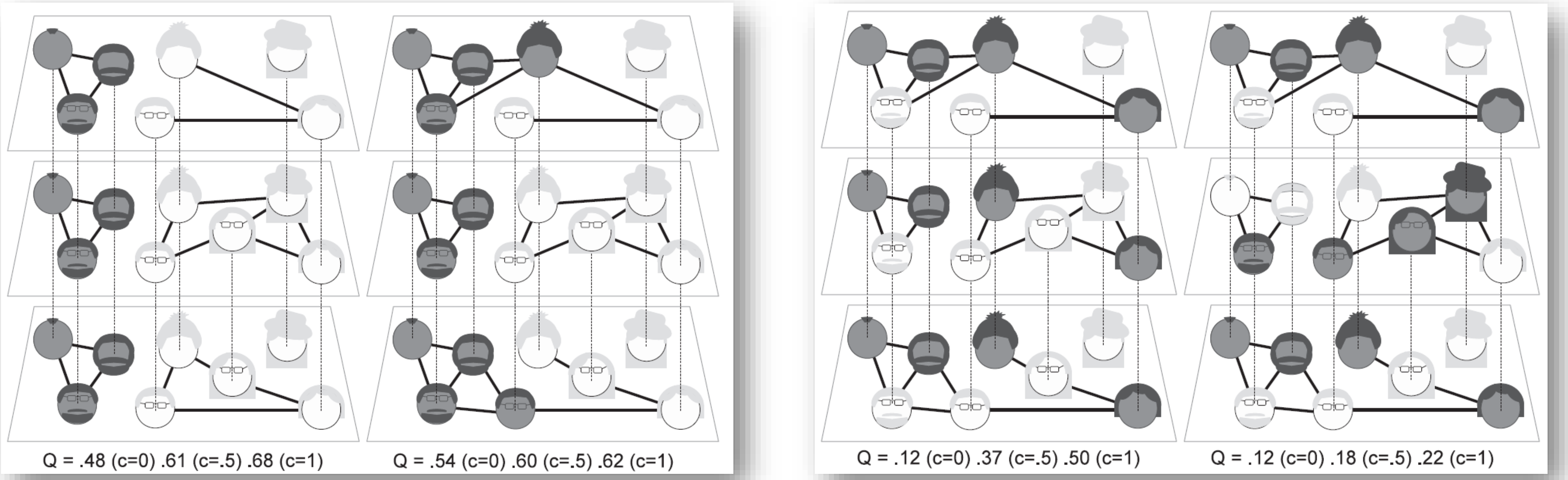
- Multilayer network modularity is defined as

$$Q_m = \frac{1}{2\mu} \sum_{ijsr} \left[\left(a_{ijs} - \frac{k_{is} k_{js}}{2m_s} \right) \delta(s, r) + c_{jsr} \delta(i, j) \right] \delta(\gamma_{i,s}, \gamma_{j,r})$$

- This extended quality function involves not just all pairs of actors (i, j) but also all pairs of layers (s, r).

Multilayer modularity of four clusterings: nodes in each layer are assigned to two clusters (black and gray); the modularity of each assignment is reported under the multilayer network using three settings:

$c_{jsr} = 1$ $\qquad$ $c_{jsr} = .5$ $\qquad$ $c_{jsr} = 0$



References

- Dickison, M. E., Magnani, M., Rossi, L. (2016). *Multilayer social networks.* Cambridge University Press. http://multilayer.it.uu.se.

- AUCS network. https://drive.google.com/drive/folders/0B_M5Zh3gg4LkNWZ5WmpoRlJhMVE

- Rossi, Luca, and Magnani, Matteo. 2015. Towards effective visual analytics on multiplex and multilayer networks. *Chaos, Solitons, and Fractals*, 72, 68–76.

- Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., Porter, M. A. (2014). *Multilayer networks.* Journal of complex networks, 2(3), 203-271. http://people.maths.ox.ac.uk/porterm/papers/multilayer_review-published.pdf.