RUSSIAN ACADEMY OF SCIENCES
INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS
(Kharkevich Institute)

IGC
Imaging Genetics Center

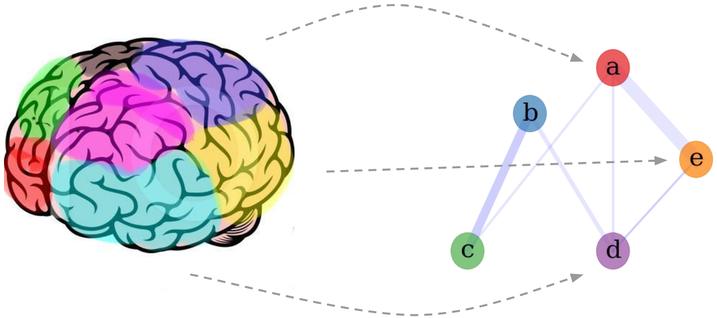
HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Classifying phenotypes based on the community structure of human brain networks

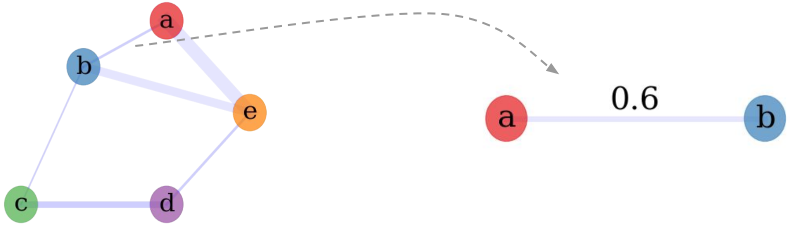
Anvar Kurmukov, Marina Ananyeva, Yulia Dodonova, Boris Gutman, Joshua Faskowitz, Neda Jahanshad, Paul Thompson, and Leonid Zhukov

Brain network

Brain regions become nodes



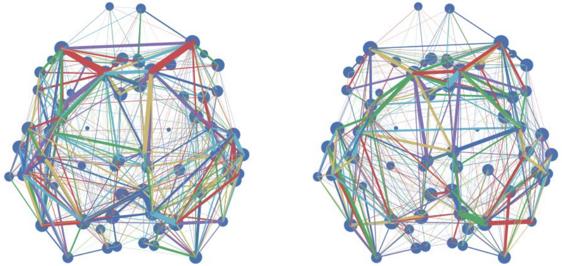
Neural connections between regions become edges



Graph $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{l}, \boldsymbol{w})$, where

- $\boldsymbol{V}$ is the set of nodes
- $\boldsymbol{E}$ is the set of edges
- $\boldsymbol{l}$ is node's labeling mapping
- $\boldsymbol{w}$ is edge's weighting mapping

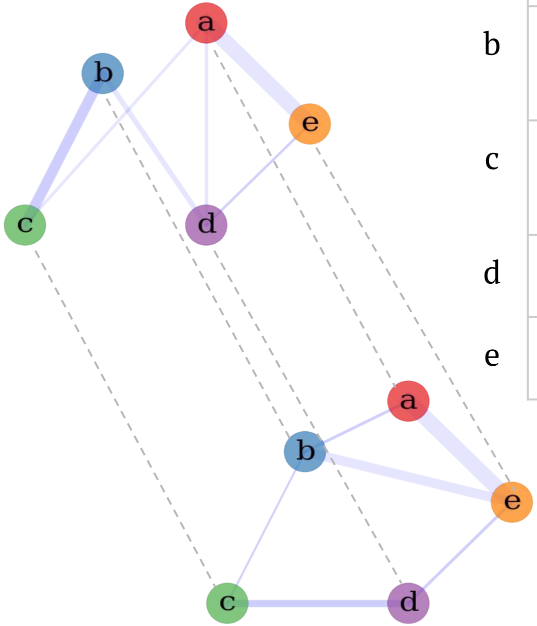
is called a brain network or a **connectome**



Connectomes: properties

- connectomes are relatively **small** graphs, usually with at most few hundreds of nodes
- the graphs are **simple** (no loops), **undirected**, i.e. the adjacency matrices are symmetric
- edges are **weighted**
- graphs are **connected**
- each node is **uniquely labeled** (according to the brain region), and the **set of labels is the same** across connectomes

Thus, graphs are fully described by their adjacency matrices (A)



	a	b	c	d	e
a	0	**6**	**0**	**0**	**30**
b	6	0	**4**	**0**	**20**
c	0	4	0	**15**	**0**
d	0	0	15	0	**6**
e	30	20	0	6	0

Adjacency matrix. Upper triangle colored

Classification task

$$\text{Dataset} = \{(G_1, y_1), (G_2, y_2), \ldots, (G_n, y_n)\}$$

G_i is simple, weighted, undirected graph with unique labels on nodes

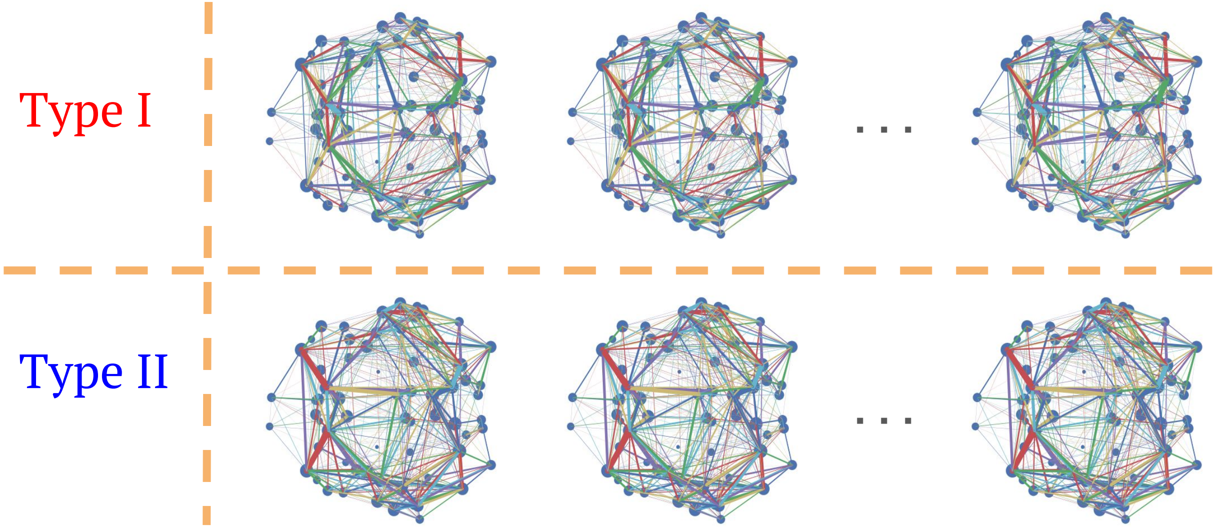
y_i is a class label

{Type I, Type II}, e.g:

{disease, control},

{male, female},

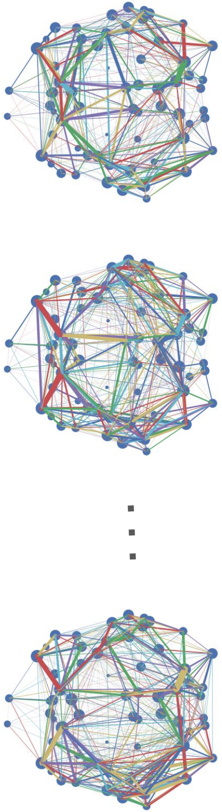
etc.



Do **Type I** graphs substantially differ from **Type II** graphs?

Kernel approach

Connectomes



⋮

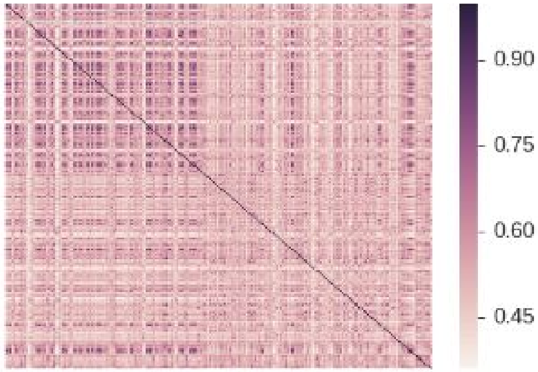
Kernel function $K(G_i, G_j)$

Kernel function $K(G_i, G_j)$

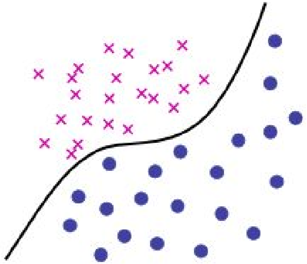
Kernel matrix

$$K : G \times G \to \mathbb{R}$$

(n samples by n samples)

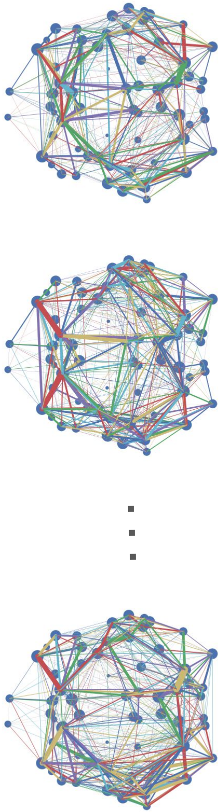


SVM with precomputed Kernel matrix



Kernel approach

Connectomes



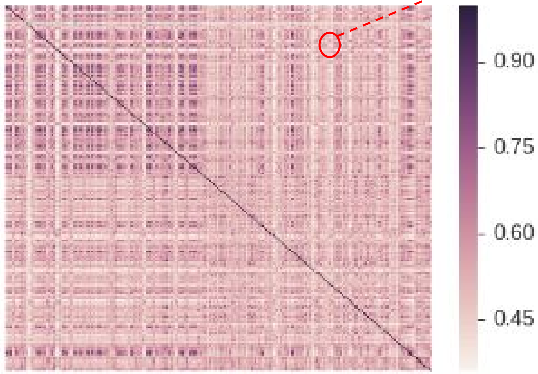
Kernel function $K(G_i, G_j)$

Kernel function $K(G_i, G_j)$

Kernel matrix

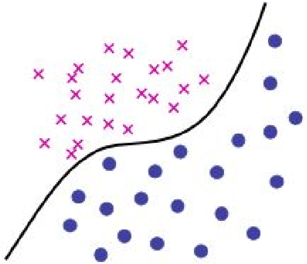
$K : G \times G \rightarrow \mathbb{R}$

(n samples by n samples)



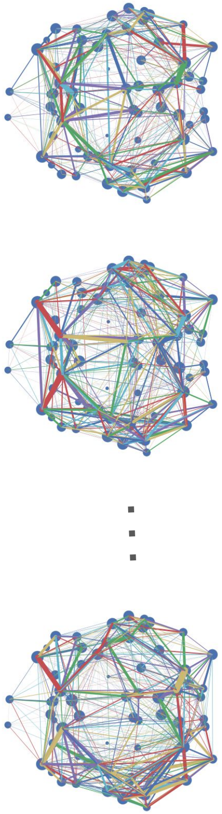
i’th **j**’th element $K(G_i, G_j)$ is a measure of similarity between G_i and G_j

SVM with precomputed Kernel matrix



Kernel approach

Connectomes



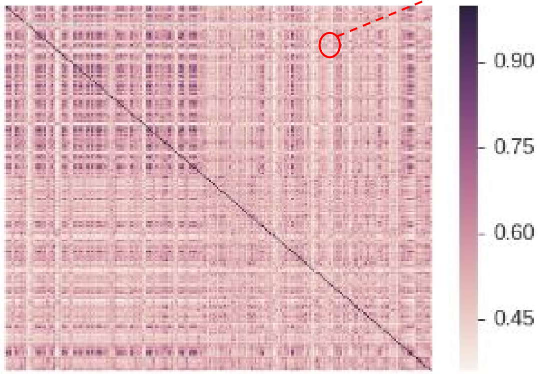
Kernel function $K(G_i, G_j)$

Kernel function $K(G_i, G_j)$

Kernel matrix

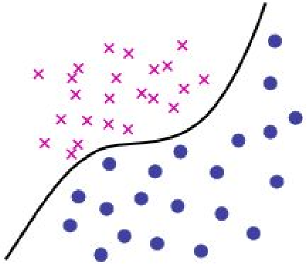
$$K : G \times G \to \mathbb{R}$$

(n samples by n samples)



i'th j'th element $K(G_i, G_j)$ is a measure of similarity between G_i and G_j

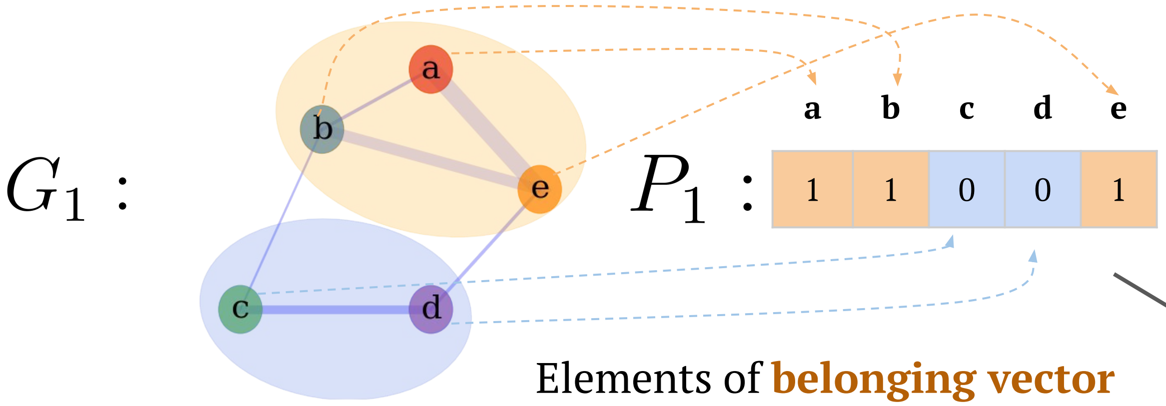
SVM with precomputed Kernel matrix



If instead of similarity measure we introduce a distance ω between the two graphs, kernel can be produced by: $K(G_i, G_j) = e^{-\alpha\omega(G_i, G_j)}$

Chan, A.B., Vasconcelos, N., Moreno, P.J. (2004) A family of probabilistic kernels based on information divergence. Univ. California, San Diego, CA, Tech. Rep. SVCL-TR-2004-1.

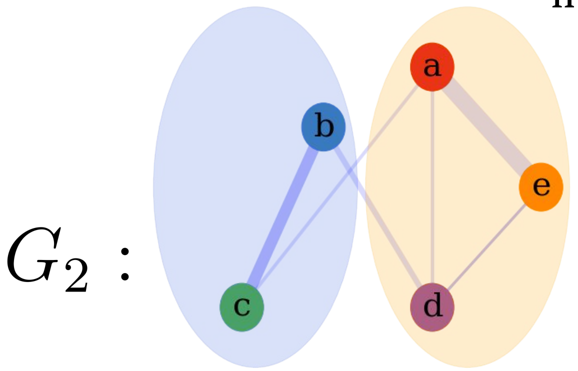
Clustering based kernels



P_1 :

a	b	c	d	e
1	1	0	0	1

Elements of **belonging vector** P_i are **belonging markers** (or indicators), vertices from the same cluster have same belonging marker



P_2 :

a	b	c	d	e
1	0	0	1	1

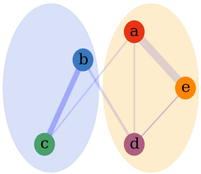
Idea: use similarity between graph clusterings as a measure of similarity between graphs

$$sim(G_1, G_2) \sim sim(P_1, P_2)$$

Note: in order to measure similarity between graph clusterings we must use appropriate metrics e.g. **Rand Index** or **Mutual Information** or their variants

Overlapping communities

Non-overlapping partition

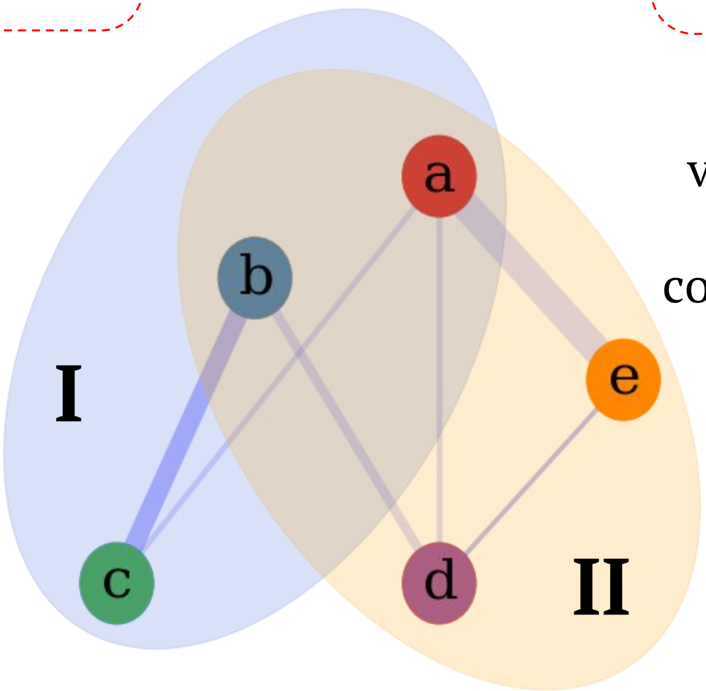


$$P_2 : \begin{array}{ccccc} \textbf{a} & \textbf{b} & \textbf{c} & \textbf{d} & \textbf{e} \\ \hline 1 & 0 & 0 & 1 & 1 \end{array}$$

Overlapping clustering

Idea: what if instead of non-overlapping partitions we will try overlapping ones?

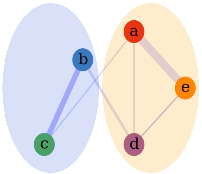
$$H_2 : \begin{array}{c|ccccc} & \textbf{a} & \textbf{b} & \textbf{c} & \textbf{d} & \textbf{e} \\ \hline \textbf{I} & .7 & .2 & 1 & 0 & 0 \\ \textbf{II} & .3 & .8 & 0 & 1 & 1 \end{array}$$



Instead of belonging vector(P_2), we now have a **belonging matrix** (H_2), columns stands for vertices, and rows stands for clusters.

Overlapping communities

Non-overlapping partition

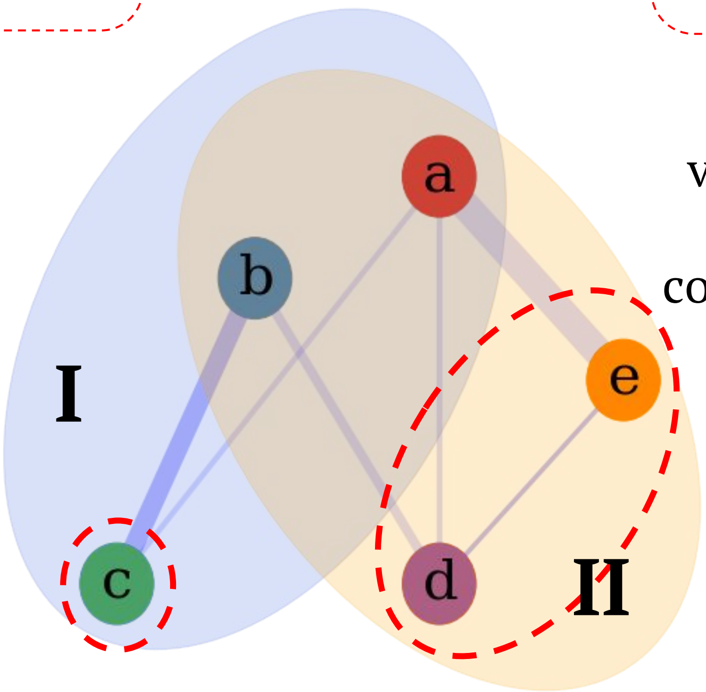


$$P_2 : \begin{array}{ccccc} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ 1 & 0 & 0 & 1 & 1 \end{array}$$

Overlapping clustering

Idea: what if instead of non-overlapping partitions we will try overlapping ones?

$$H_2 : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & .3 & .8 & 1 & 0 & 0 \\ \mathbf{II} & .7 & .2 & 0 & 1 & 1 \end{array}$$

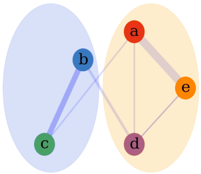


Instead of belonging vector(P_2), we now have a **belonging matrix** (H_2), columns stands for vertices, and rows stands for clusters.

Nodes ***c, d, e,*** belong to one cluster (***c*** to **I** and ***d, e*** to **II**)

Overlapping communities

Non-overlapping partition

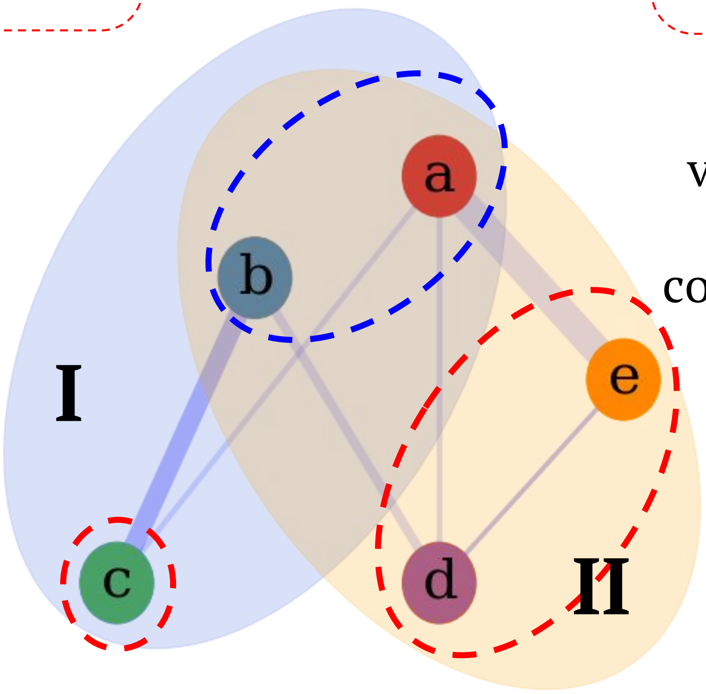


$$P_2 : \begin{array}{ccccc} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ 1 & 0 & 0 & 1 & 1 \end{array}$$

$$H_2 : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & .3 & .8 & 1 & 0 & 0 \\ \mathbf{II} & .7 & .2 & 0 & 1 & 1 \end{array}$$

Nodes ***a*** and ***b*** belong to both clusters with **different strength**

Overlapping clustering



Idea: what if instead of non-overlapping partitions we will try overlapping ones?

Instead of belonging vector(P_2), we now have a **belonging matrix** (H_2), columns stands for vertices, and rows stands for clusters.

Nodes ***c, d, e,*** belong to one cluster (***c*** to **I** and ***d, e*** to **II**)

Similarity of graph clusterings

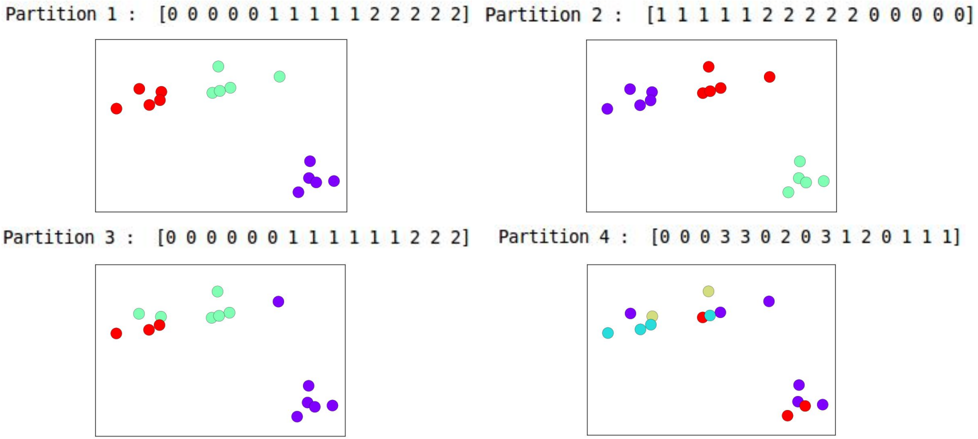
AMI (P_1, P_1) = 1.0

AMI (P_1, P_2) = 1.0

AMI (P_1, P_3) = 0.529

AMI (P_1, P_4) = 0.049

Mutual Information is a measure of similarity, thus value 1 indicates completely identical partitions, and values close to 0 stands for very dissimilar. Which is also true for both AMI and NMI



```
Partition 1 :  [0 0 0 0 0 1 1 1 1 1 2 2 2 2 2]
Partition 2 :  [1 1 1 1 1 2 2 2 2 2 0 0 0 0 0]
Partition 3 :  [0 0 0 0 0 0 1 1 1 1 1 1 2 2 2]
Partition 4 :  [0 0 0 3 3 0 2 0 3 1 2 0 1 1 1]
```

Vinh, N.X., Epps, J., Bailey, J.: Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. Journal of Machine Learning Research, pp. 2837–2854 (2010)

In order to compare non-overlapping partitions we used so-called **Adjusted Mutual Information**

$$\omega_{AMI}(G_i, G_j) = 1 - AMI(P_i, P_j)$$

All **soft** overlappings were thresholded down to **hard** overlappings

Normalized Mutual Information were used to compare hard overlappings

$$\omega_{NMI}(G_i, G_j) = 1 - NMI(\overline{H_i}, \overline{H_j})$$

where overlined H's are hard overlappings, produced from soft ones

$$K(G_i, G_j) = e^{-\alpha\omega(G_i, G_j)}$$

Soft and Hard overlapping

SOFT belonging matrix

$$H_2 : \begin{array}{c|c|c|c|c|c|} & \textbf{a} & \textbf{b} & \textbf{c} & \textbf{d} & \textbf{e} \\ \hline \textbf{I} & .7 & .2 & 1 & 0 & 0 \\ \hline \textbf{II} & .3 & .8 & 0 & 1 & 1 \\ \hline \end{array}$$

HARD belonging matrix

Threshold = 0.15

Threshold = 0.4

$$\overline{H_2} : \begin{array}{c|c|c|c|c|c|} & \textbf{a} & \textbf{b} & \textbf{c} & \textbf{d} & \textbf{e} \\ \hline \textbf{I} & 1 & 1 & 1 & 0 & 0 \\ \hline \textbf{II} & 1 & 1 & 0 & 1 & 1 \\ \hline \end{array} \qquad \begin{array}{c|c|c|c|c|c|} & \textbf{a} & \textbf{b} & \textbf{c} & \textbf{d} & \textbf{e} \\ \hline \textbf{I} & 1 & 0 & 1 & 0 & 0 \\ \hline \textbf{II} & 0 & 1 & 0 & 1 & 1 \\ \hline \end{array}$$

Soft and Hard overlapping

SOFT belonging matrix

$$H_2 : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & .7 & .2 & 1 & 0 & 0 \\ \mathbf{II} & .3 & .8 & 0 & 1 & 1 \end{array}$$

Strength / weights
(e.g. **a** **belongs** to **I** with strength .7
and to **II** with strength .3)

From strength to
cluster indicators

HARD belonging matrix

Threshold = 0.15

$$\overline{H_2} : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & 1 & 1 & 1 & 0 & 0 \\ \mathbf{II} & 1 & 1 & 0 & 1 & 1 \end{array}$$

Threshold = 0.4

$$\begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & 1 & 0 & 1 & 0 & 0 \\ \mathbf{II} & 0 & 1 & 0 & 1 & 1 \end{array}$$

Cluster
indicators
(e.g. **a** **lies** in **I**
and **II**)

Soft and Hard overlapping

SOFT belonging matrix

$$H_2 : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & .7 & .2 & 1 & 0 & 0 \\ \mathbf{II} & .3 & .8 & 0 & 1 & 1 \end{array}$$

Strength / weights
(e.g. **a** **belongs** to **I** with strength .7
and to **II** with strength .3)

From strength to cluster indicators

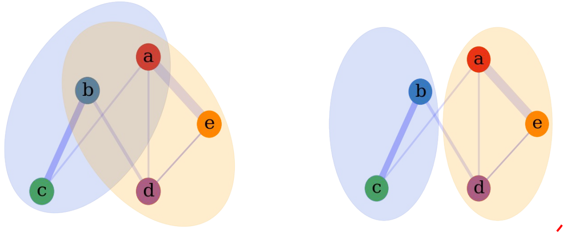
HARD belonging matrix

Threshold = 0.15 ⇢ Threshold = 0.4

$$\overline{H_2} : \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & 1 & 1 & 1 & 0 & 0 \\ \mathbf{II} & 1 & 1 & 0 & 1 & 1 \end{array} \qquad \begin{array}{c|ccccc} & \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} & \mathbf{e} \\ \hline \mathbf{I} & 1 & 0 & 1 & 0 & 0 \\ \mathbf{II} & 0 & 1 & 0 & 1 & 1 \end{array}$$

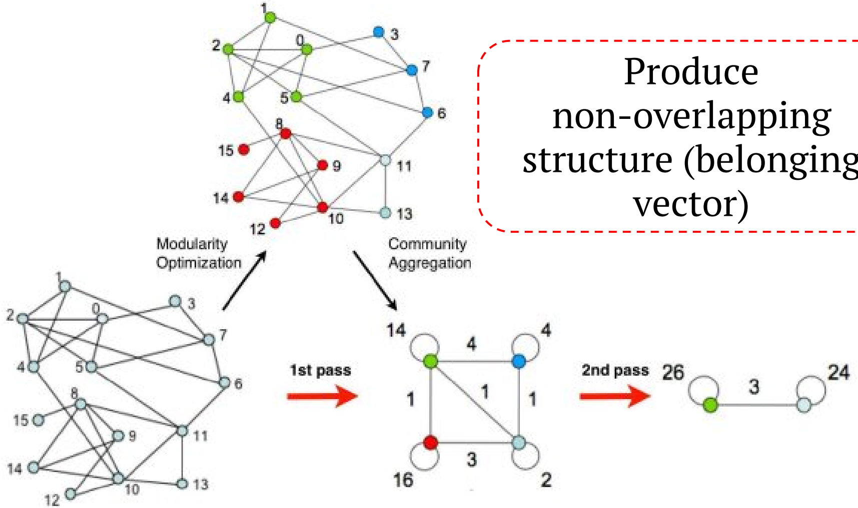
Cluster indicators
(e.g. **a** **lies** in **I**
and **II**)

Increasing threshold leads to non-overlapping structure



Graph clustering methods

Louvain Modularity



Produce non-overlapping structure (belonging vector)

Modularity is given by:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(i,j)$$

Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008.10 (2008): P10008.

Non-negative Matrix Factorization (NMF)

$$\min_{W,H \geq 0} ||A - WH||_F^2$$

$$A \in \mathbb{R}_+^{n \times n}, W \in \mathbb{R}_+^{n \times k}, H \in \mathbb{R}_+^{k \times n}$$

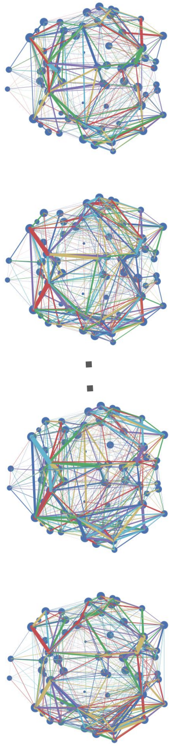
H is interpreted as a **belonging matrix**.

Produce overlapping community structure (belonging matrix)

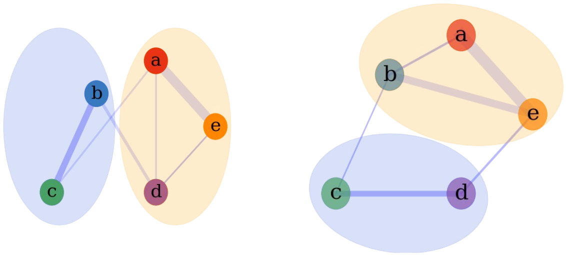
Kuang, Da, Chris Ding, and Haesun Park. "Symmetric nonnegative matrix factorization for graph clustering." *Proceedings of the 2012 SIAM international conference on data mining*. Society for Industrial and Applied Mathematics, 2012.

Summary

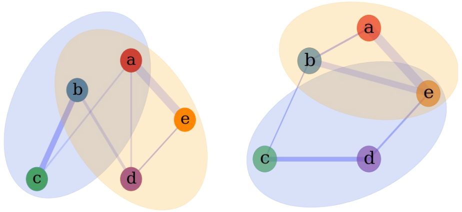
Connectomes



Non-overlapping partitions

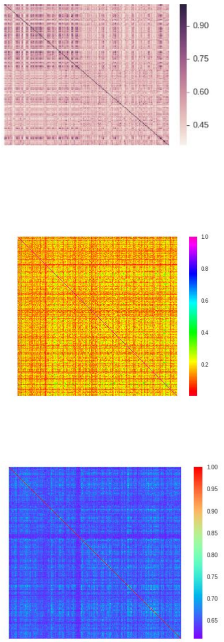


Overlapping clusters

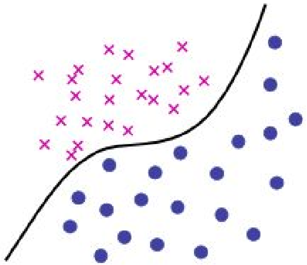


Frobenius norm between adjacency matrices (all pairwise comparisons)

Kernel matrix



SVM with precomputed Kernel matrix



Data description

We use the Alzheimer‘s Disease Neuroimaging Initiative (ADNI2) database which comprises a total of 228 individuals (756 scans). Each individual has at least 1 brain scan and at most 6 scans.

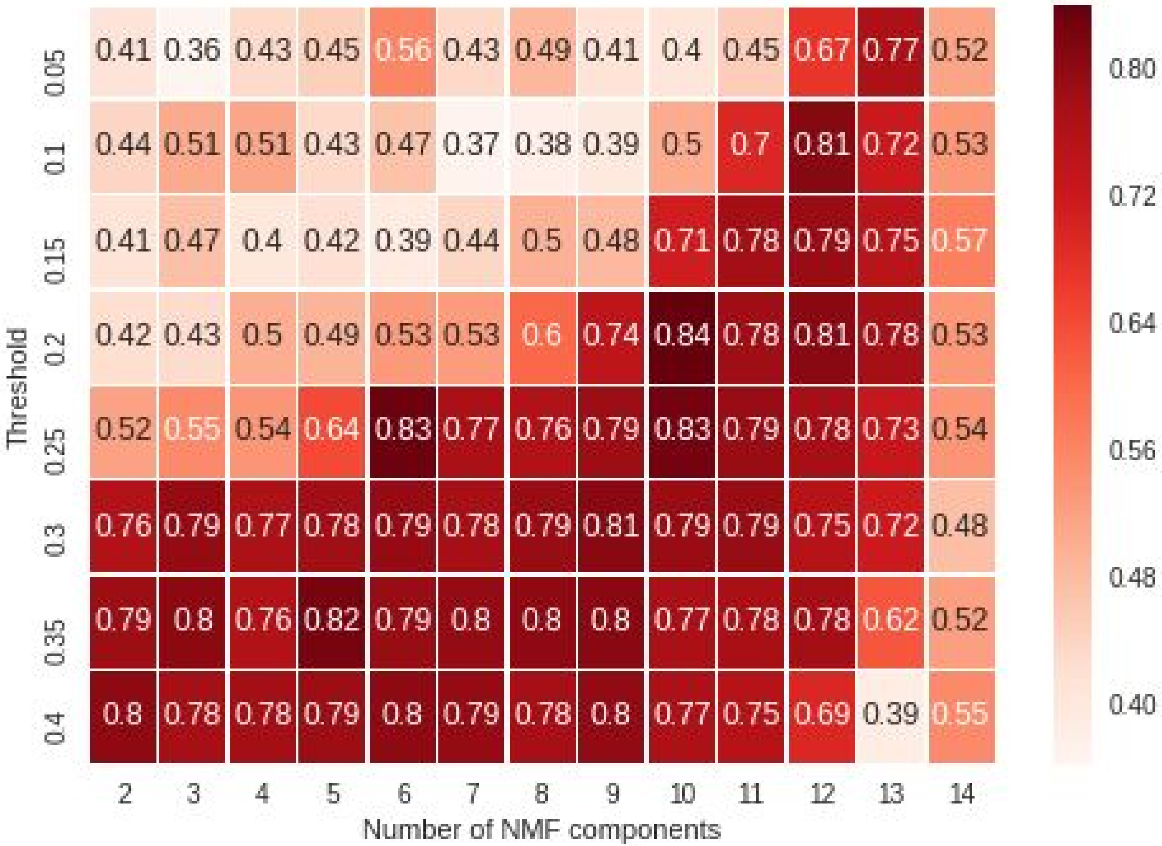
The data include 47 people with **AD** (136 AD scans), 40 individuals with **LMCI** (147 LMCI scans), 80 individuals with **EMCI** (283 EMCI scans), and 61 **healthy (NC)** participants (190 scans).

Thus we consider four classification tasks:

- AD vs NC
- AD vs LMCI
- LMCI vs EMCI
- EMCI vs NC

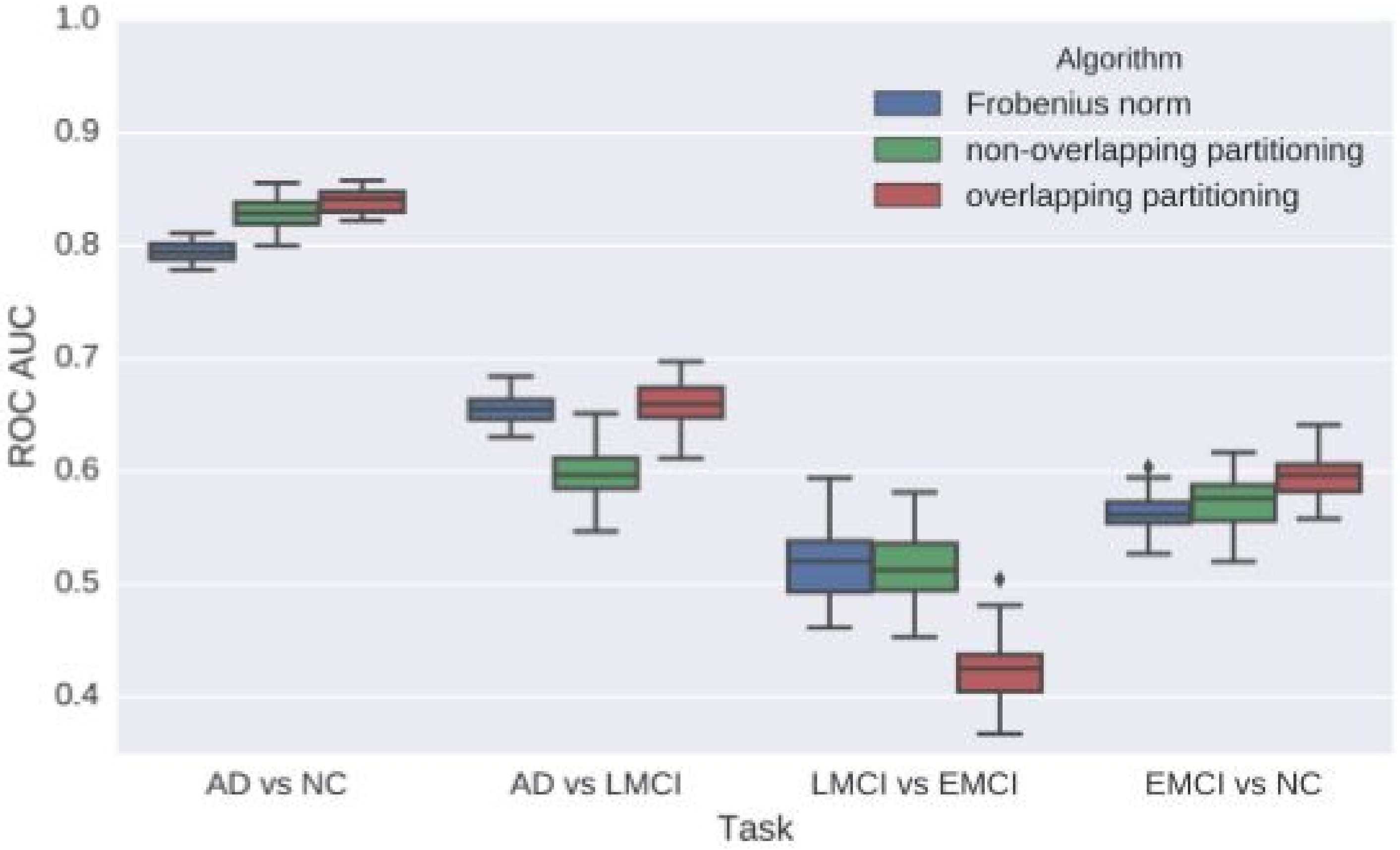
All ROC AUCs were obtained on 10-fold subject based cross validation and averaged over 50 different randomizations. Model parameters (number of NMF components, SVM penalty, and kernel parameter) were tuned on another 50 different randomization

Classification results. NMF.



Threshold \ Number of NMF components	2	3	4	5	6	7	8	9	10	11	12	13	14
0.05	0.41	0.36	0.43	0.45	0.56	0.43	0.49	0.41	0.4	0.45	0.67	0.77	0.52
0.1	0.44	0.51	0.51	0.43	0.47	0.37	0.38	0.39	0.5	0.7	0.81	0.72	0.53
0.15	0.41	0.47	0.4	0.42	0.39	0.44	0.5	0.48	0.71	0.78	0.79	0.75	0.57
0.2	0.42	0.43	0.5	0.49	0.53	0.53	0.6	0.74	0.84	0.78	0.81	0.78	0.53
0.25	0.52	0.55	0.54	0.64	0.83	0.77	0.76	0.79	0.83	0.79	0.78	0.73	0.54
0.3	0.76	0.79	0.77	0.78	0.79	0.78	0.79	0.81	0.79	0.79	0.75	0.72	0.48
0.35	0.79	0.8	0.76	0.82	0.79	0.8	0.8	0.8	0.77	0.78	0.78	0.62	0.52
0.4	0.8	0.78	0.78	0.79	0.8	0.79	0.78	0.8	0.77	0.75	0.69	0.39	0.55

Classification results. Overview.



Conclusion

Does community structure of human brain networks provide enough information to classify phenotypes?

We proposed a framework to **compare** both overlapping and non-overlapping community structures of brain networks within the machine learning settings

We demonstrated the performance of the proposed pipeline in a task of **classifying Alzheimer‘s disease**, mild cognitive impairment, and healthy participants

Models that made full use of **overlapping community structures** performed slightly better than those based on non-overlapping community structures

Thank you for your attention!

Classifying phenotypes based on the community structure of human brain networks

Anvar Kurmukov
kurmukovai@gmail.com

Unauthorized figures

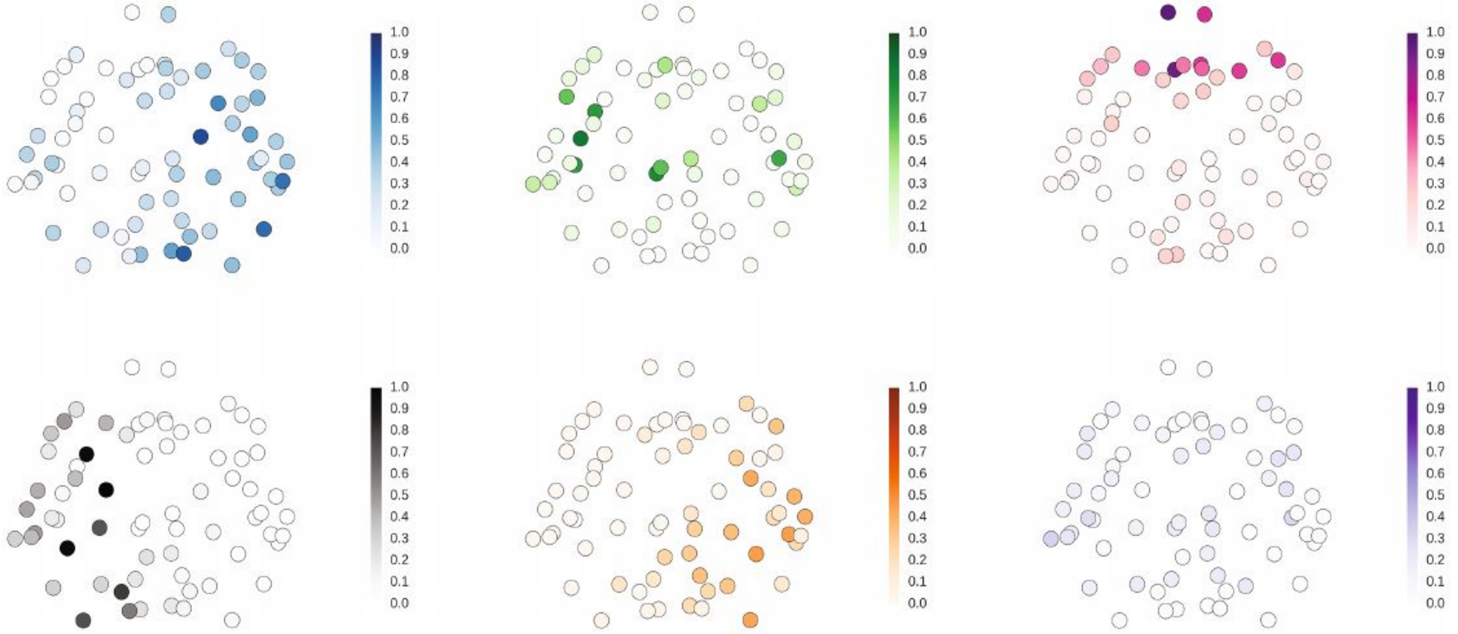


Fig. 2. Six overlapping communities: an example of a single network (healthy subject) with the nodes shown in their original 3D coordinates (axial view); color intensity is proportional to the strength of belonging to the respective community

Unauthorized figures

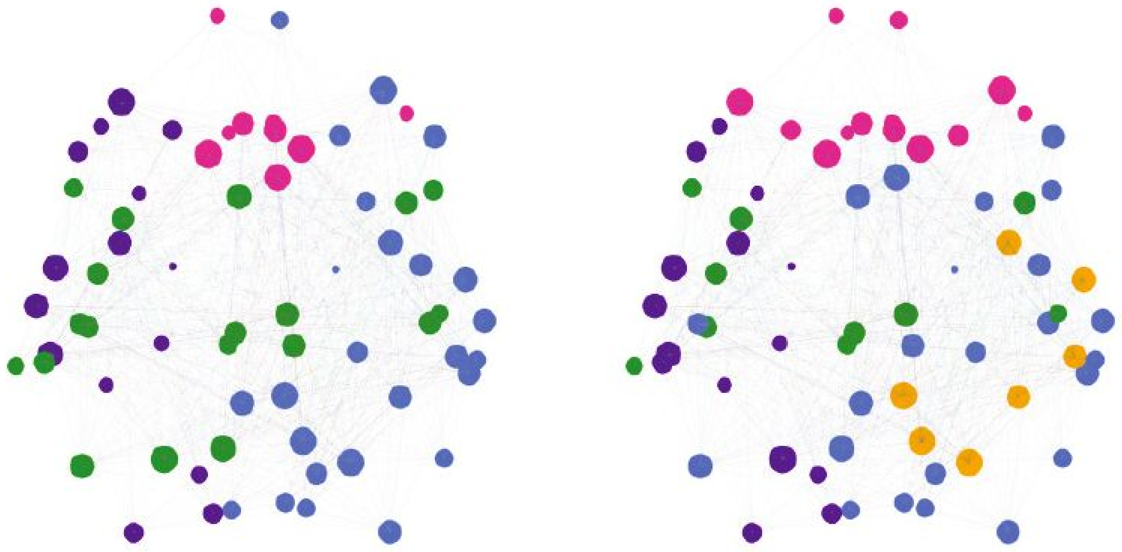


Fig. 3. Comparison of the non-overlapping (left) and overlapping (right) community structures obtained for the same example graph as in Fig. 3; node size is proportional to its degree (the number of edges coming from the respective node). Right plot is produced by selecting a single community for each node based on the maximal membership probability.