

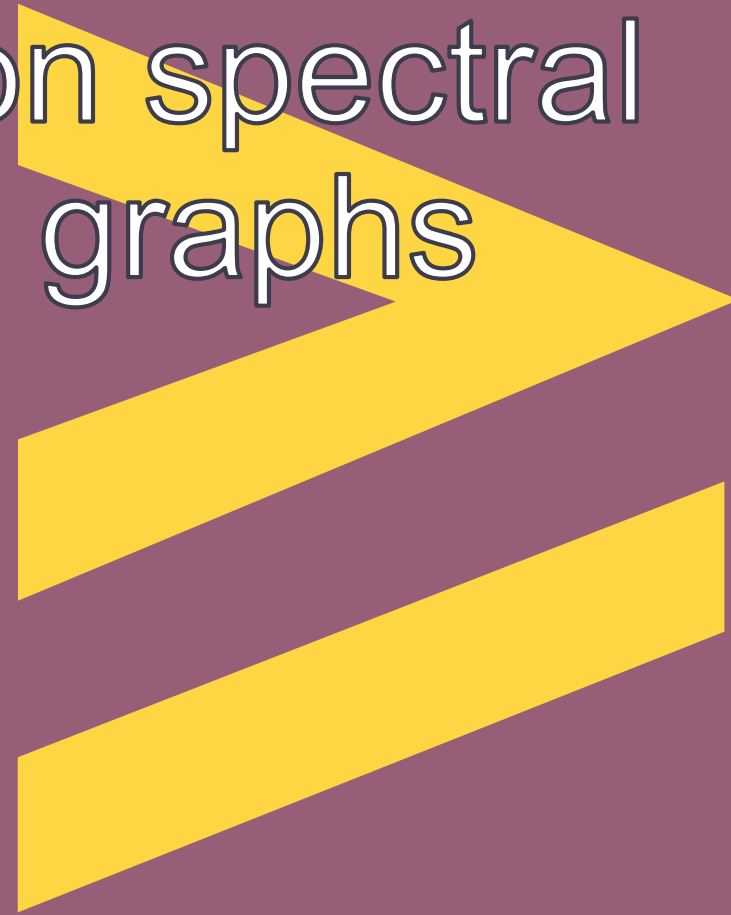


Wydział  
Fizyki

POLITECHNIKA WARSZAWSKA

# Overlapping communities detection based on spectral analysis of line graphs

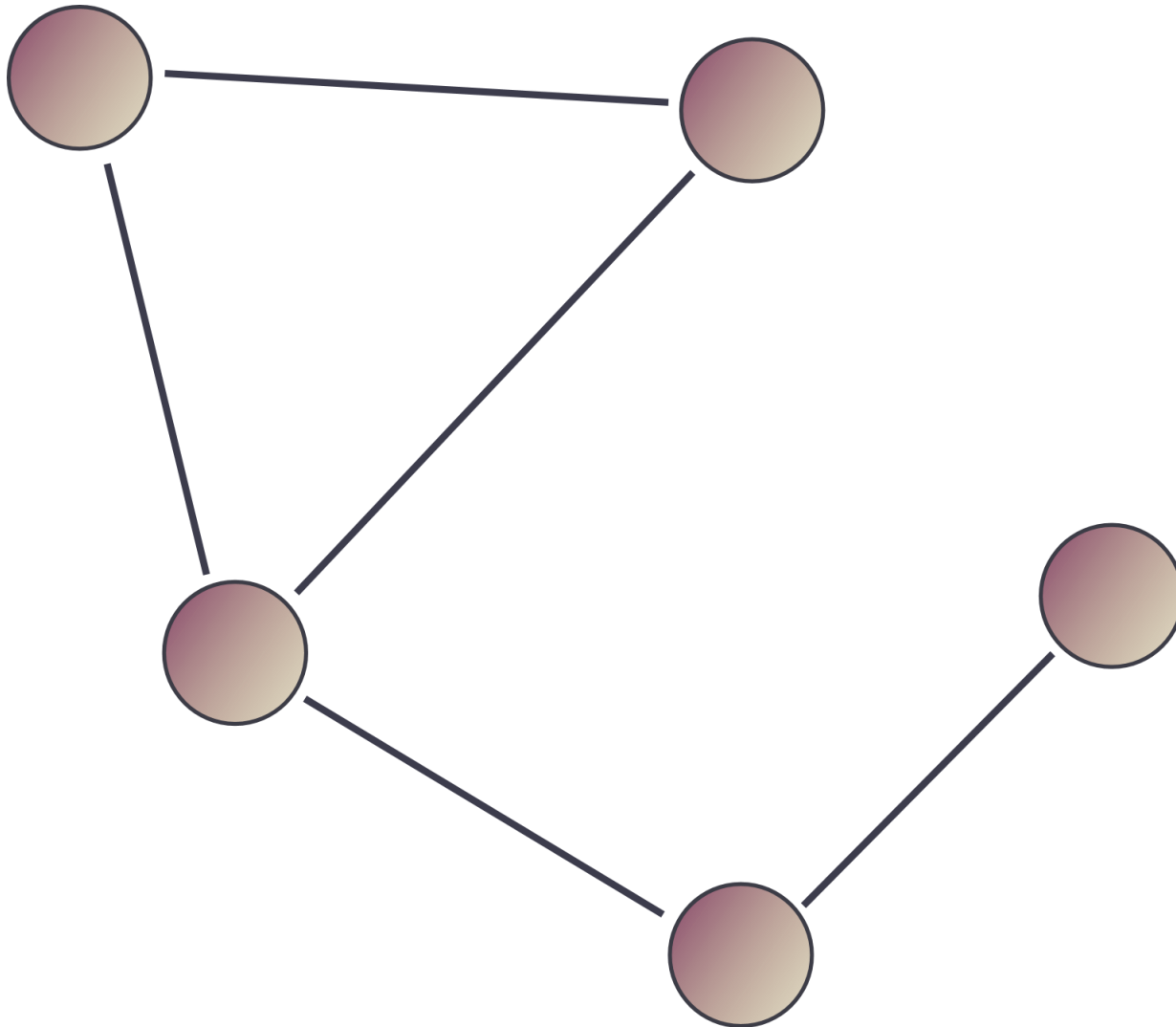
Patryk A. Bojarski



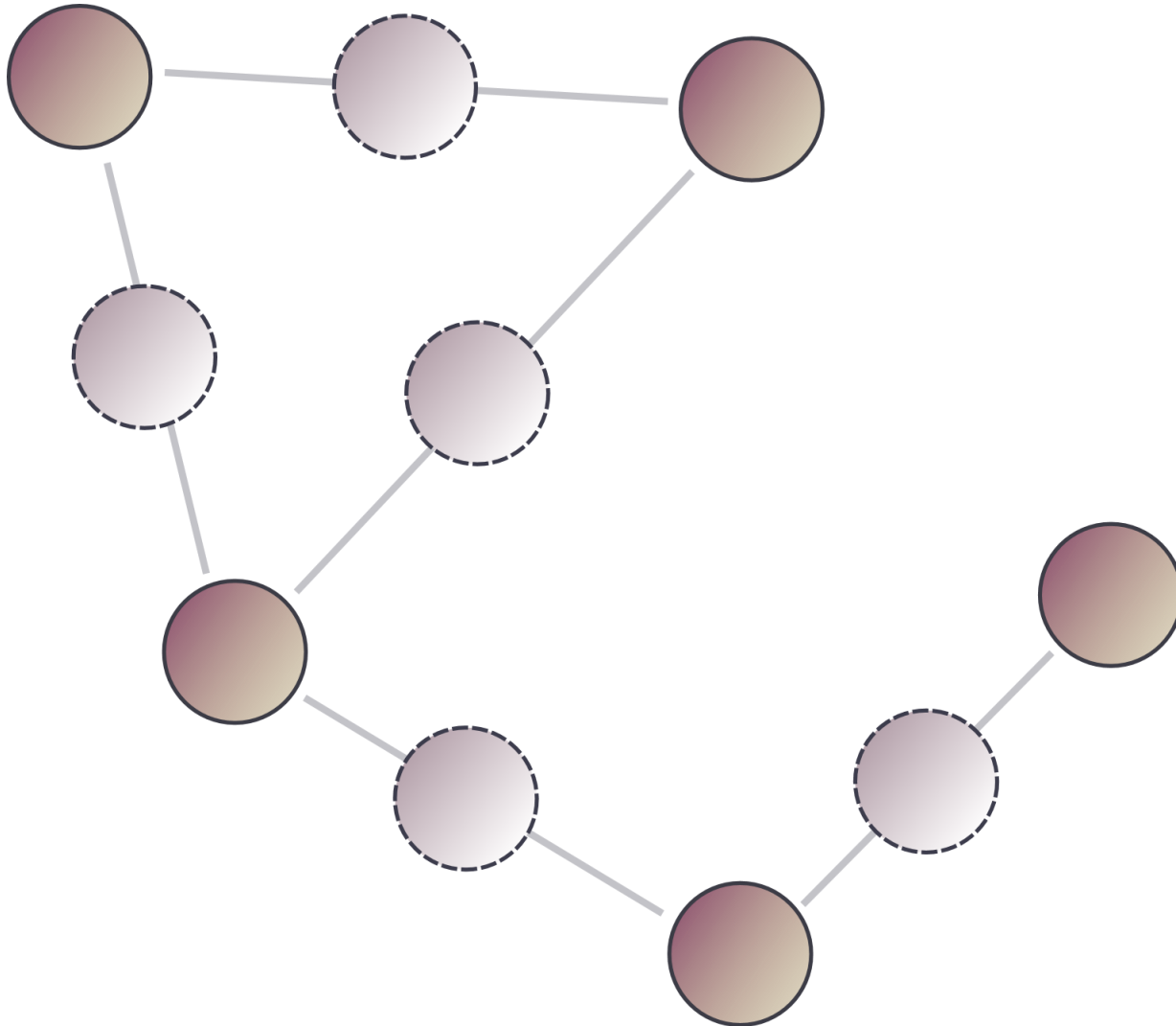


# Line graphs

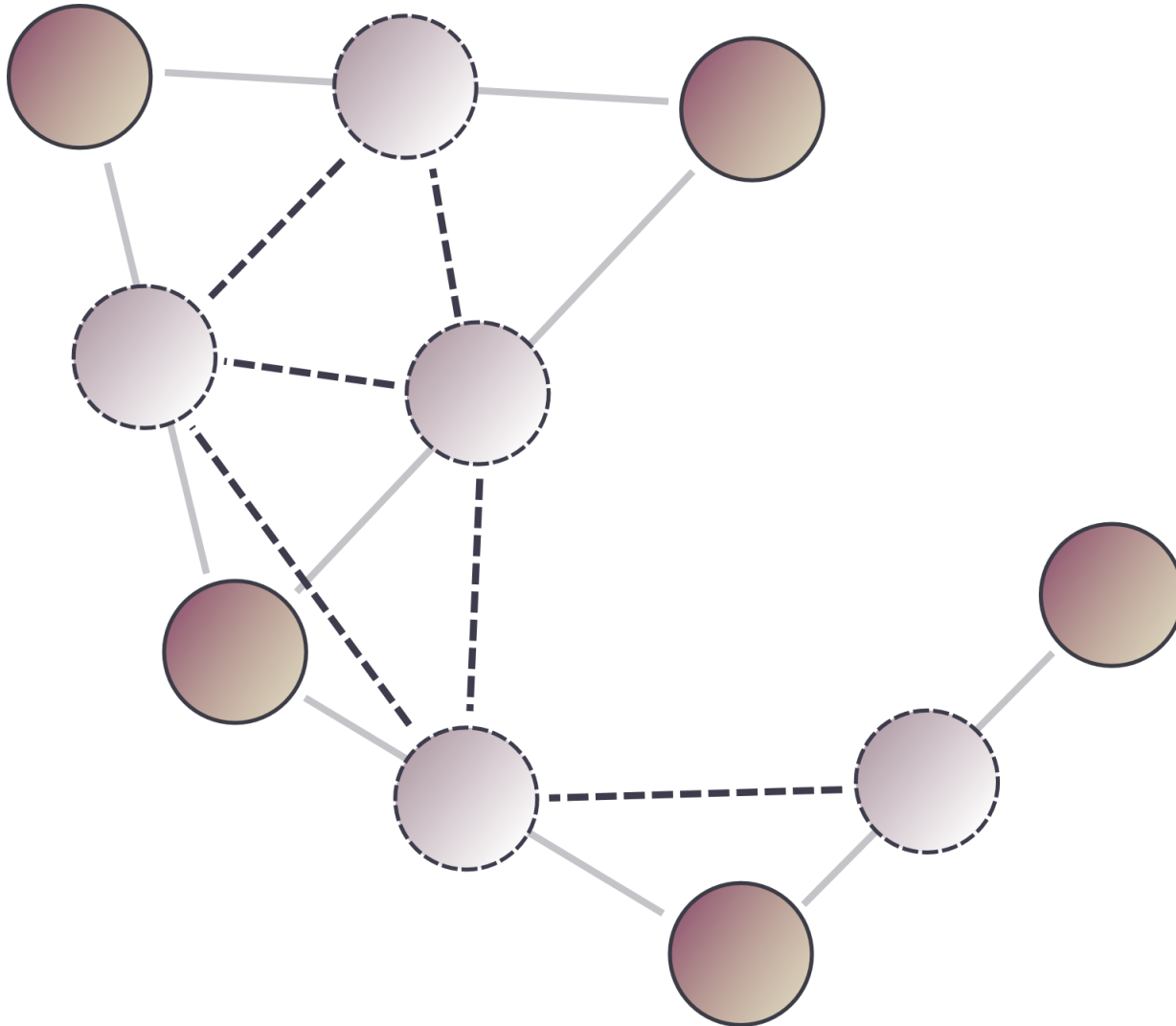
# Line graph – example



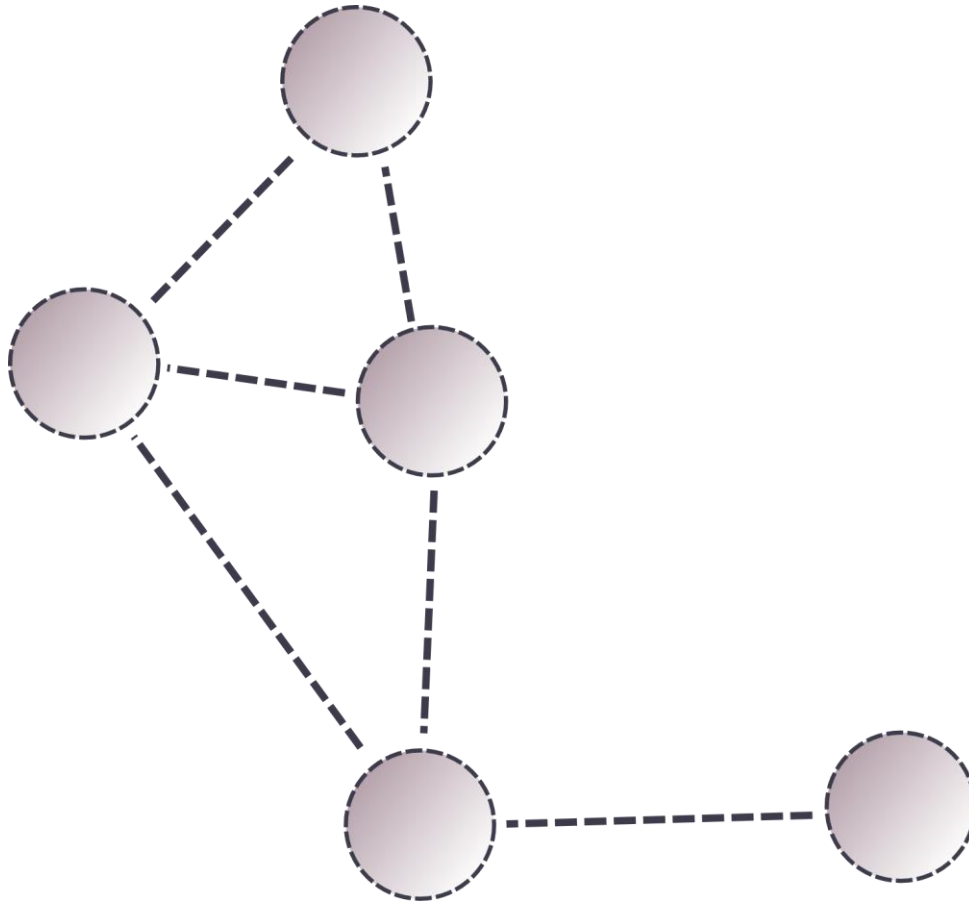
# Line graph – example



# Line graph – example



# Line graph – example





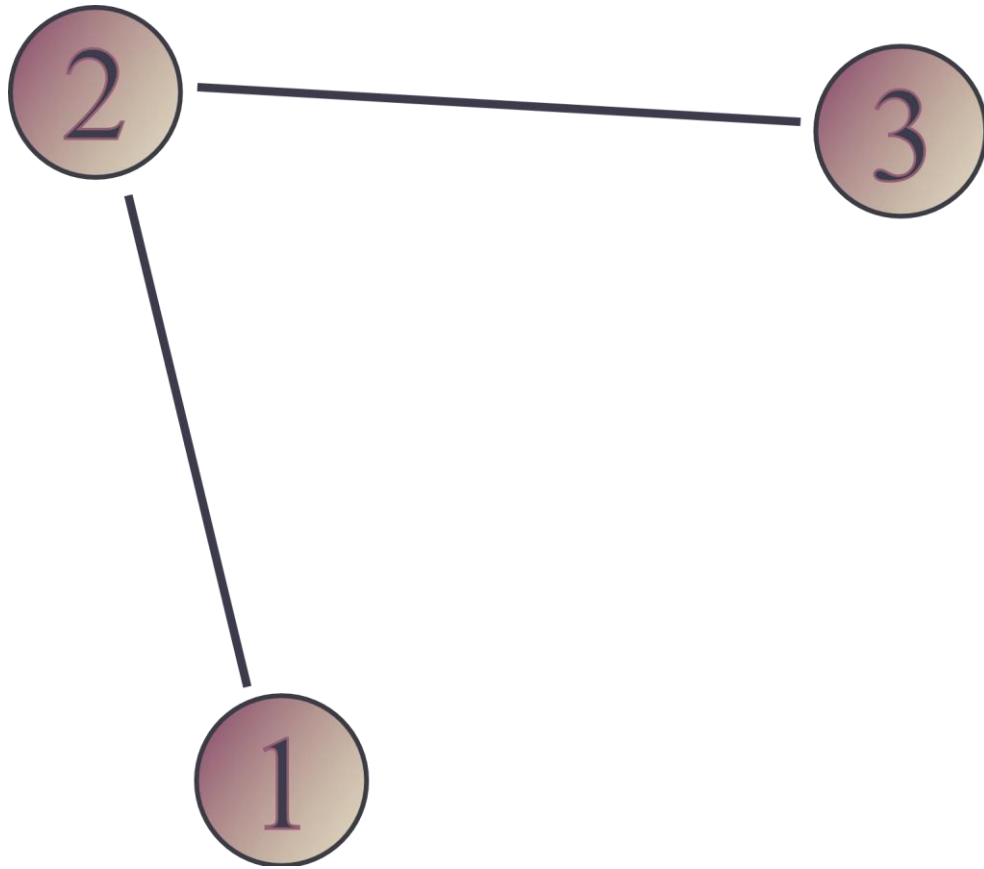
Angular distance

$$d_{AB}^A = \arccos \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| \cdot |\vec{B}|}$$





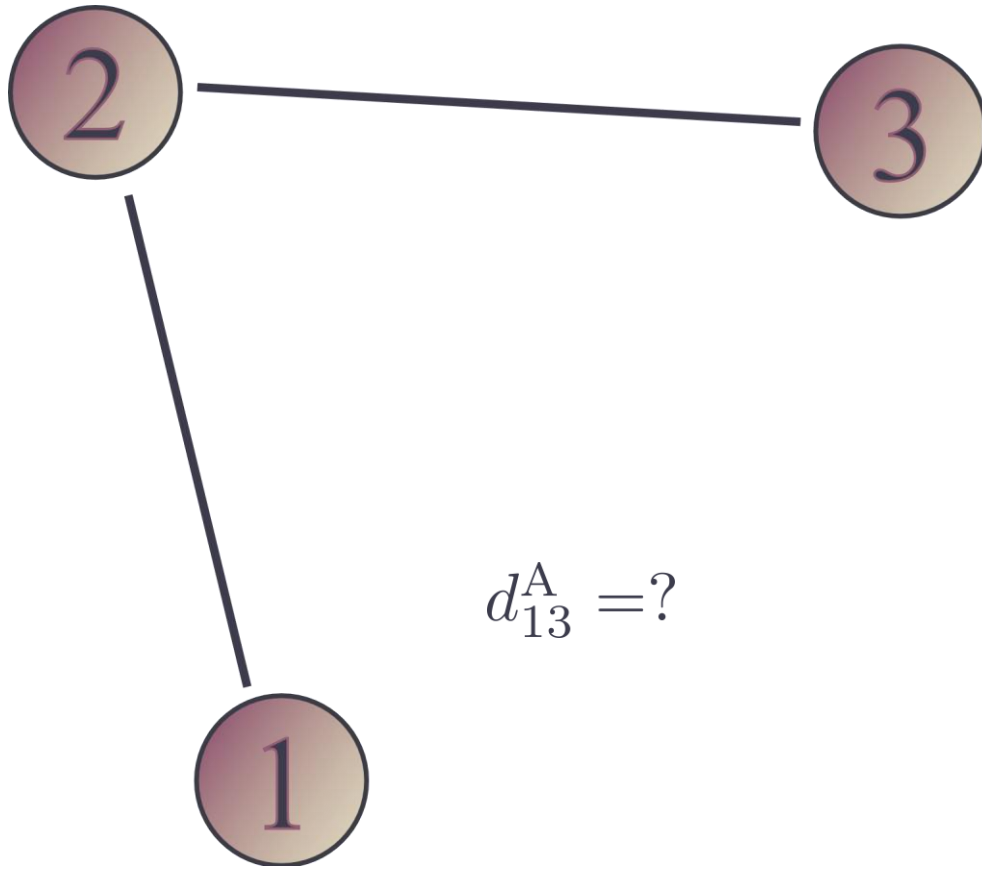
$$d_{AB}^A = \arccos \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| \cdot |\vec{B}|}$$



$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



$$d_{AB}^A = \arccos \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| \cdot |\vec{B}|}$$

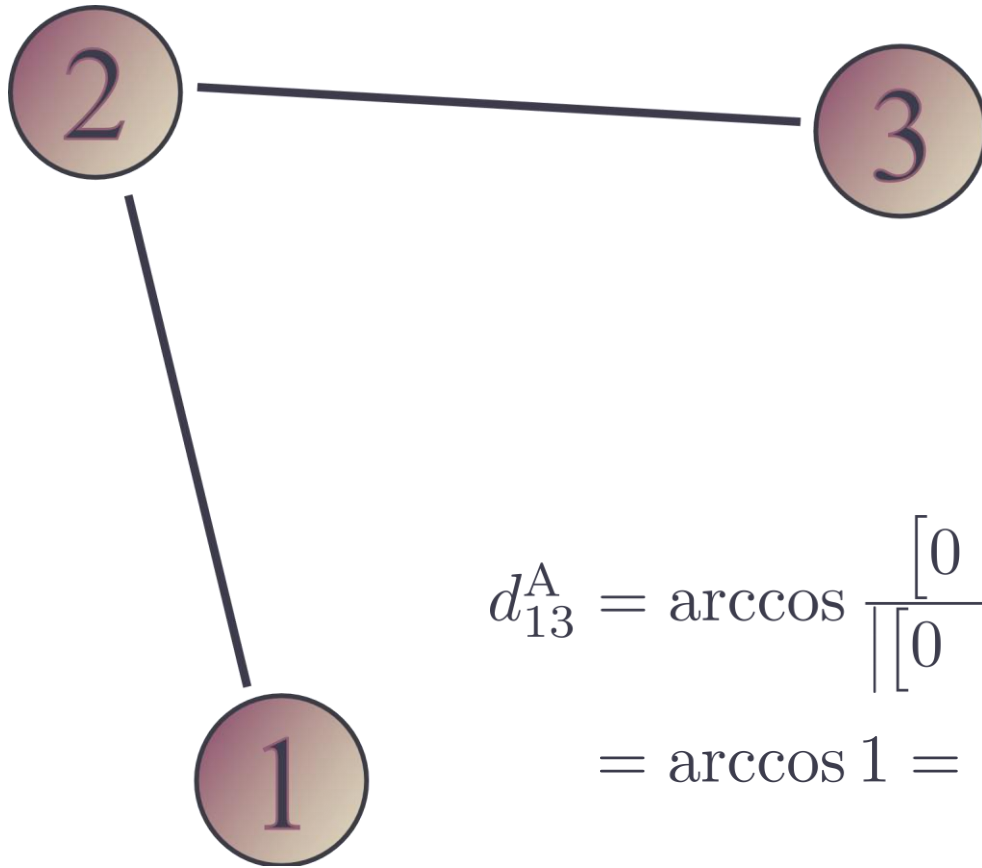


$$d_{13}^A = ?$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$



$$d_{AB}^A = \arccos \frac{\vec{A} \cdot \vec{B}}{|\vec{A}| \cdot |\vec{B}|}$$



$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\begin{aligned} d_{13}^A &= \arccos \frac{[0 \ 1 \ 0] \cdot [0 \ 1 \ 0]}{\|[0 \ 1 \ 0]\| \cdot \|[0 \ 1 \ 0]\|} \\ &= \arccos 1 = 0 \end{aligned}$$





The partition density



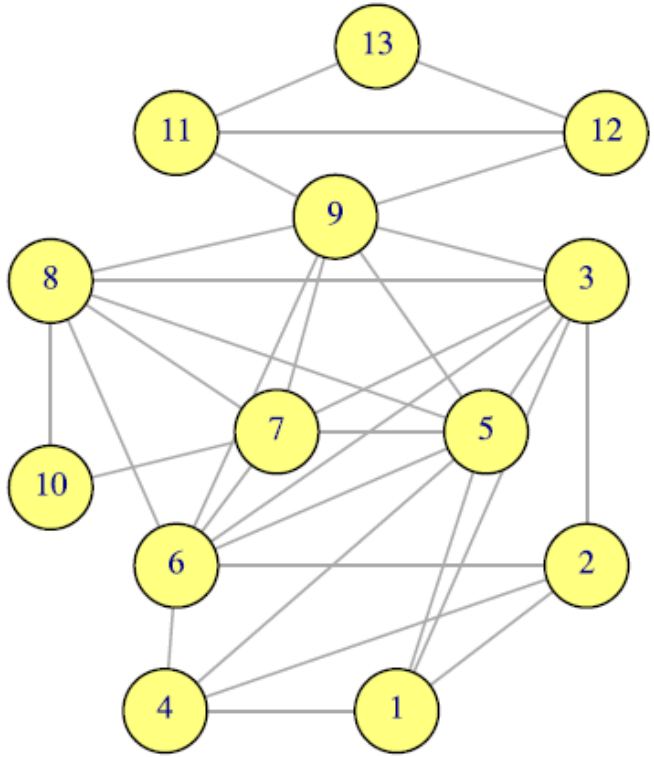
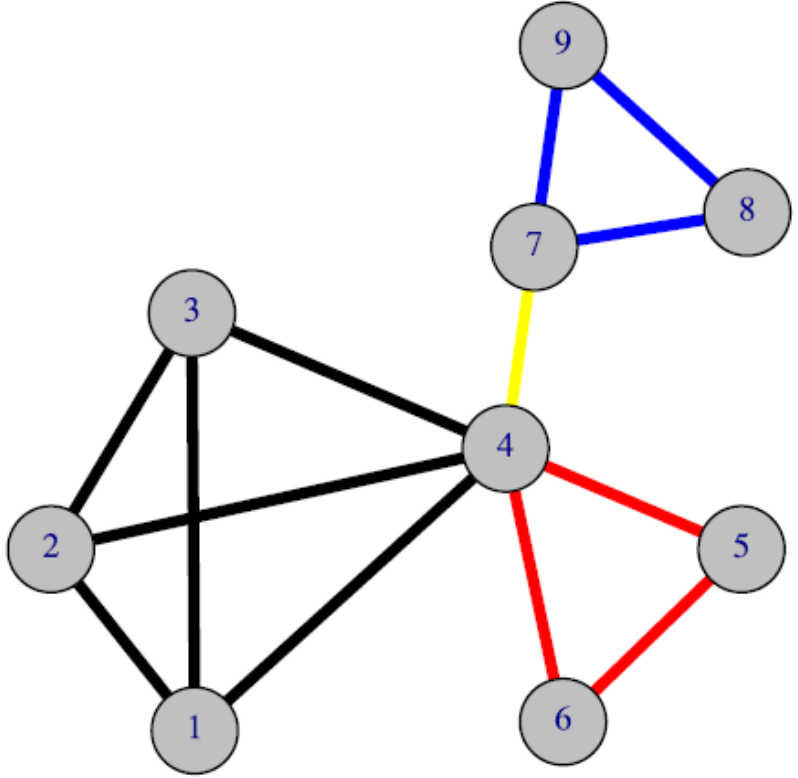
# Definition

$$D_{XZ} = \frac{2}{m} \sum_c \frac{m_c}{q_c} \frac{m_c - (n_c - 1)}{(n_c - 2) \cdot (n_c - 1)}$$

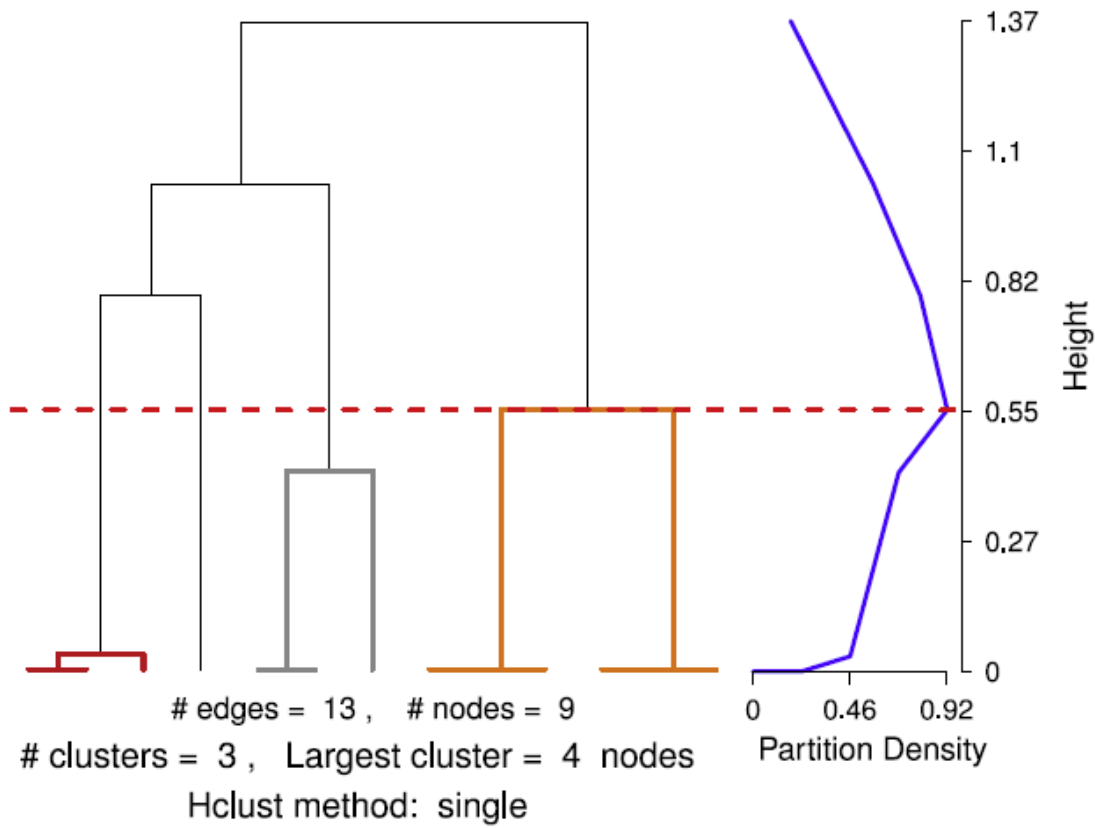
$$q_c = \max_{j \in c} I_j$$



# Results



# Link Communities Dendrogram



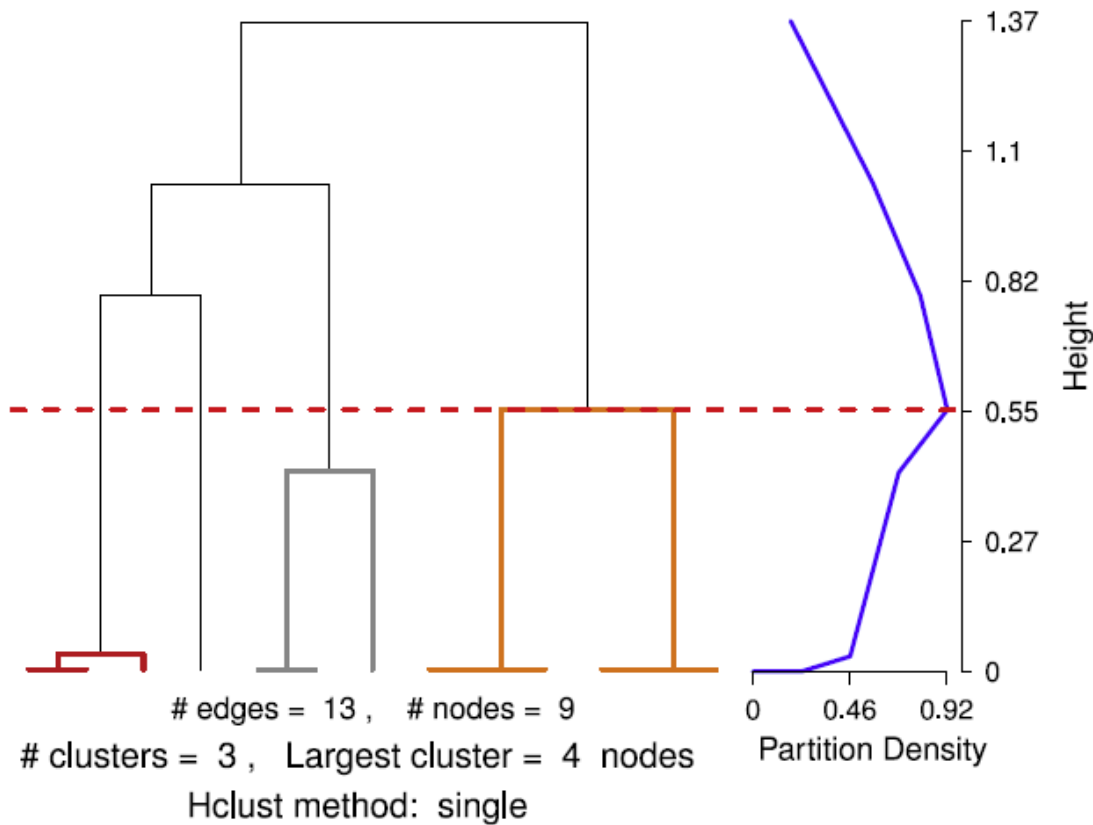
## Community Membership

	1	2	3	$\Sigma$
4				2
7				1
8				1
9				1
1				1
2				1
3				1
5				1
6				1
$\Sigma$	3	4	3	





## Link Communities Dendrogram



## Community Membership

	1	2	3	$\Sigma$
4				2
7				1
8				1
9				1
1				1
2				1
3				1
5				1
6				1
$\Sigma$	3	4	3	

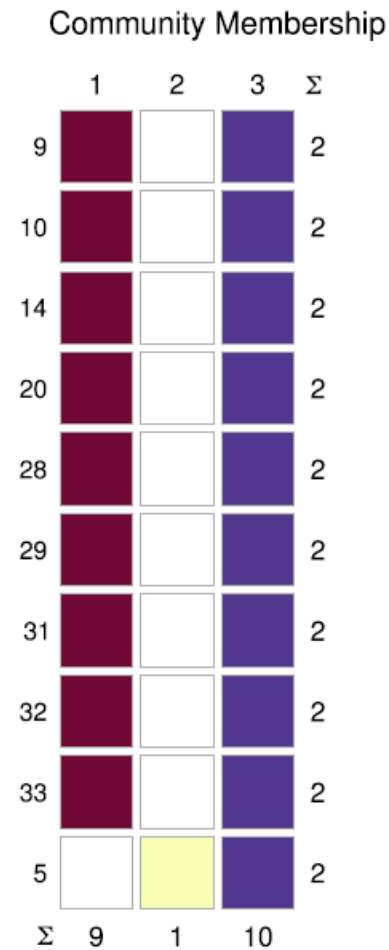
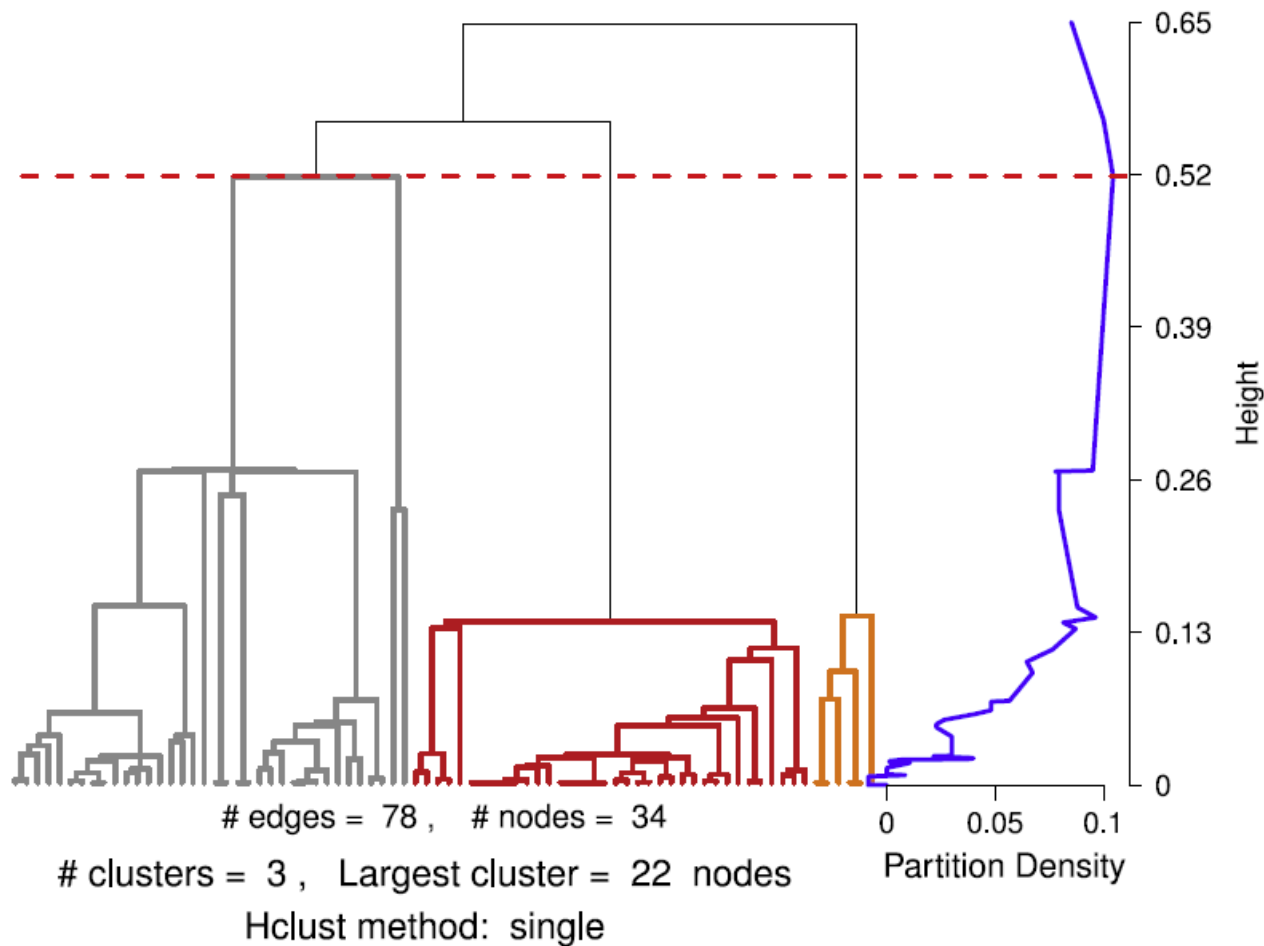
The results of four algorithms on Standard network.

Metrics	SAoLG	Ahn's algorithm	CPM	GN
LM	0.4645	0.4645	0.4645	0.3412
PD	0.9231	0.9231	0.9231	0.6923
CN	3	3	3	3
CR	100%	100%	100%	100%
UV	0	0	0	0

Note: LM means modularity of line graph, PD means partition density, CN means community number, CR means coverage rate, UV means uncovered vertices.



# Link Communities Dendrogram



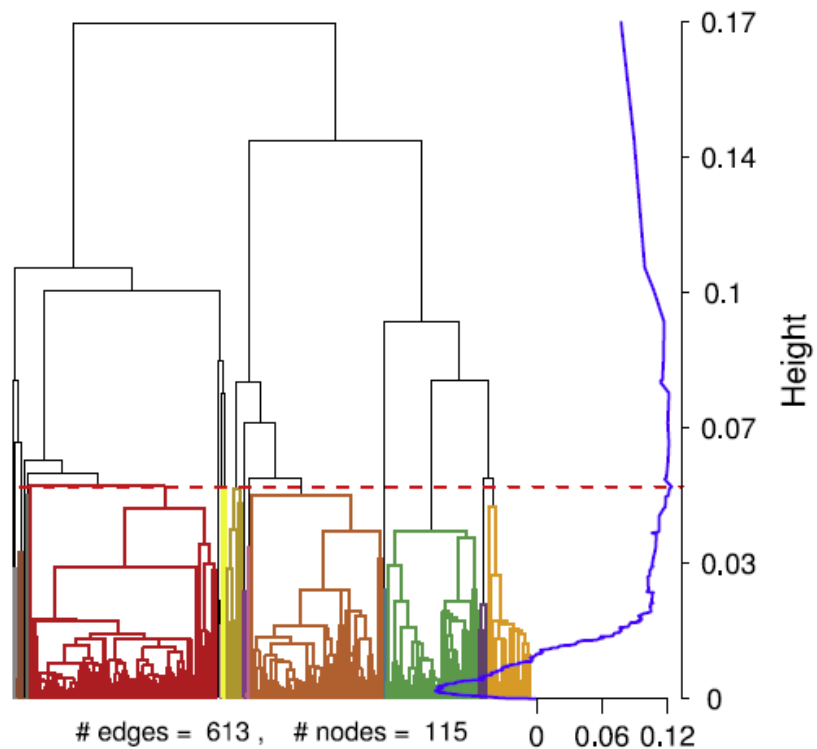
The results of four algorithms on Karate network.

Metrics	SAoLG	Ahn's algorithm	CPM	GN
LM	0.48	<b>0.68*</b>	0.23	0.40
PD	0.10	<b>0.28*</b>	0.20	-0.019
CN	<b>3*</b>	8	<b>3*</b>	5
CR	<b>100%*</b>	97.06%	94.12%	97.06%
UV	<b>0*</b>	1	2	1

Note: LM means modularity of line graph, PD means partition density, CN means community number, CR means coverage rate, UV means uncovered vertices.

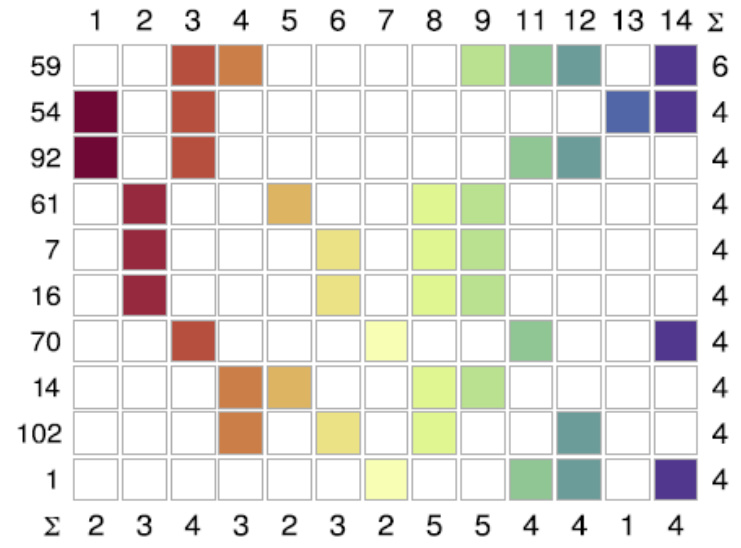


Link Communities Dendrogram



# edges = 613 , # nodes = 115  
 # clusters = 14 , Largest cluster = 49 nodes  
 Hclust method: single

Community Membership



The results of four algorithms on Football network.

Metrics	SAoLG	Ahn's algorithm	CPM	GN
LM	<b>0.89*</b>	0.18	0.28	0.6
PD	0.12	<b>0.55*</b>	0.54	0.4642567
CN	14	26	<b>13*</b>	10
CR	<b>100%*</b>	96.5%	98%	<b>100%</b>
UV	<b>0*</b>	4	2	<b>0</b>





# Bibliography

- [1] **Chun Gui, Ruisheng Zhang, Rongjing Hu, Gouming Huang, Jiaxuan Wei**  
*Overlapping communities detection based on spectral analysis of line graphs*  
Physica A 498 (2018) 50-65
- [2] **Dr hab. Pior Fronczak**  
*Algebra grafów*  
<http://if.pw.edu.pl/~agatka/sieci/algebra.pdf>



Thanks for your  
attention!