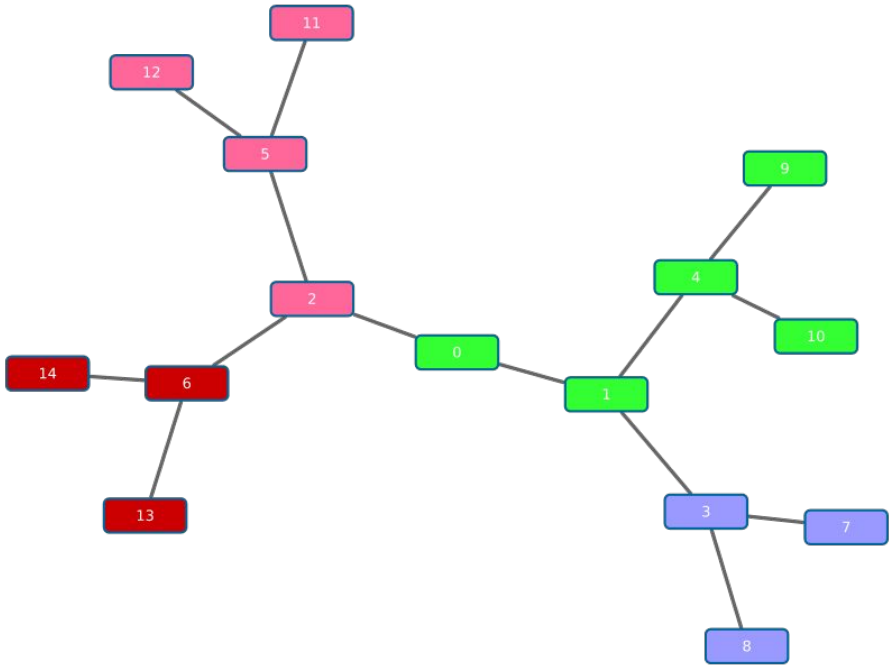
A Flexible Fitness Function for

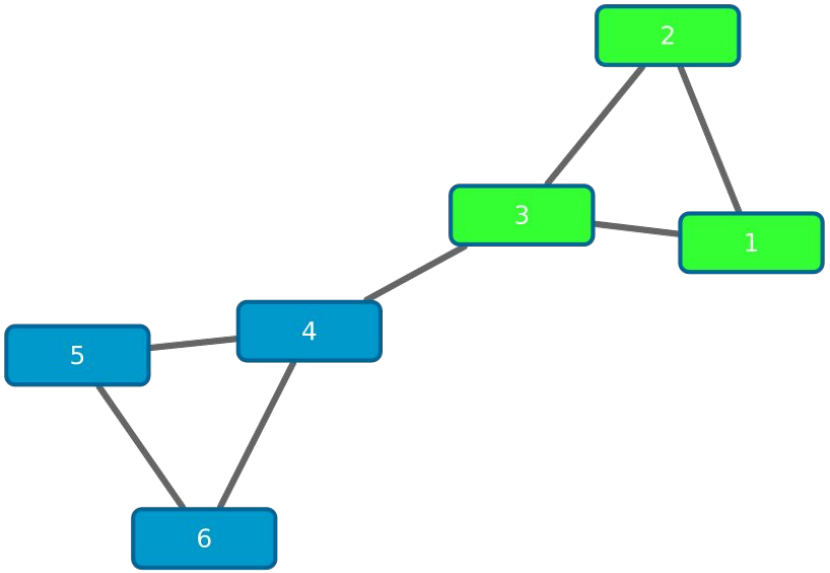
Community Detection In Complex Networks

Fabrício Olivetti de França and Guilherme Palermo Coelho
Universidade Federal do ABC / Universidade Estadual de Campinas

What is a community?

Many definitions!





How to find a community

One way to find a community is to formulate it as an optimization problem.

Use any suitable optimization algorithm available.

Modularity

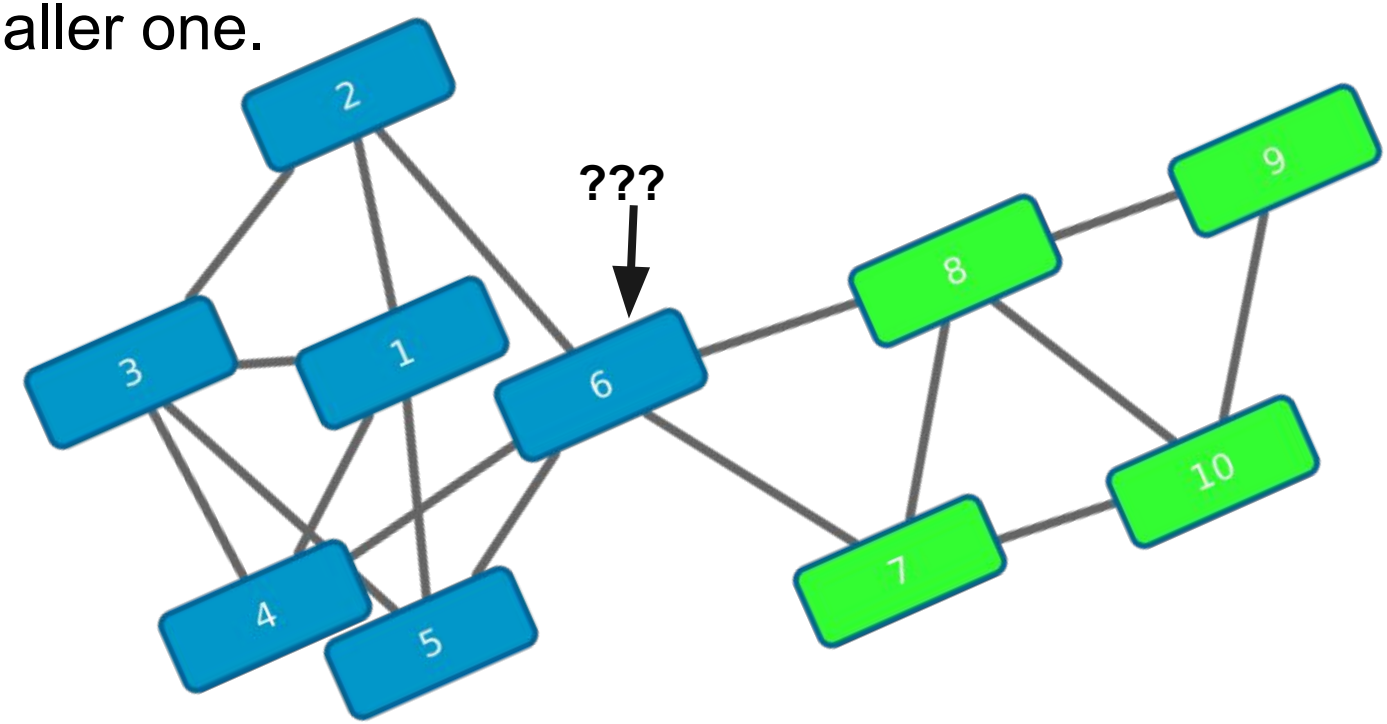
A common objective-function is the Modularity:

$$Q = \sum_i (e_{i,j} - (\sum_j e_{i,j})^2)$$

where $e_{i,j}$ is the fraction of edges connecting community i to community j.

Modularity

If a node has 3 edges connecting it to community *i* and 2 edges to a smaller community *j*, it will be assigned to the larger community, regardless of its contribution to the smaller one.



FLEX: A Flexible Fitness Function

Let's define FLEX for a given partitioning **p**:

$$FLEX(p) = \frac{1}{|V|} \sum_{c \in p} CC(c)$$

where CC(c) is the contribution of community **c**.

FLEX

The community contribution is calculated as:

$$CC(c) = \sum_{i \in c} LC(i, c) - \frac{|c|^{\gamma}}{|V|}$$

where:

LC(i,c) is the contribution of node i to community c

γ is the importance of having smaller communities.

FLEX

LC(i,c) is defined as:

$$LC(i,c) = \alpha * \triangle(i,c) + (1-\alpha) * N(i,c) - \beta * \wedge(i,c)$$

$\triangle$(i,c): transitivity of i on c by transitivity of i

N(i,c): neighbors of i in c by neighbors of i

$\wedge$(i,c): # of open triangles with i in c by # of open triangles with i

FLEX

LC(i,c) is defined as:

$$LC(i,c) = \alpha * \triangle(i,c) + (1-\alpha) * N(i,c) - \beta * \wedge(i,c)$$

α: importance of clustering x neighborhood

β: penalization for open triangles

FLEX

High transitivity x High connectivity

Penalizes open triangles (if desired) and very large communities.

Overlapping Nodes

An insight:

if $\Delta(i,c) < thr\Delta$ or $N(i,c) < thrN$:

Add i to $\forall c'$ where $N(i,c') > thrSh$

Experiments

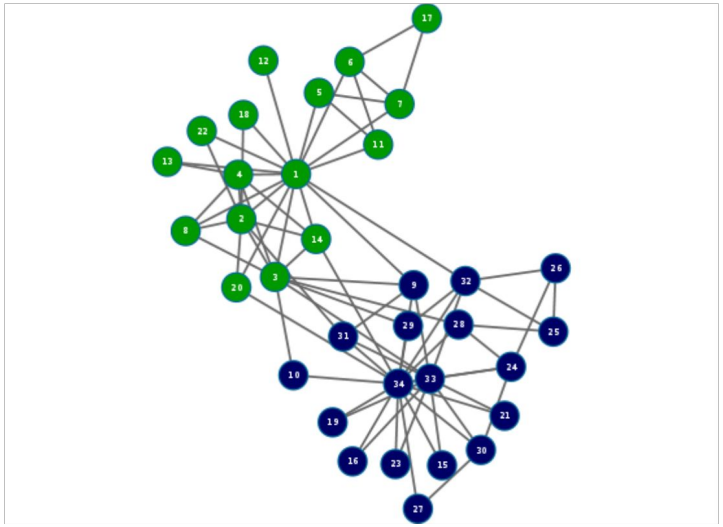
Four real networks.

Four randomly generated without noise and overlapping nodes.

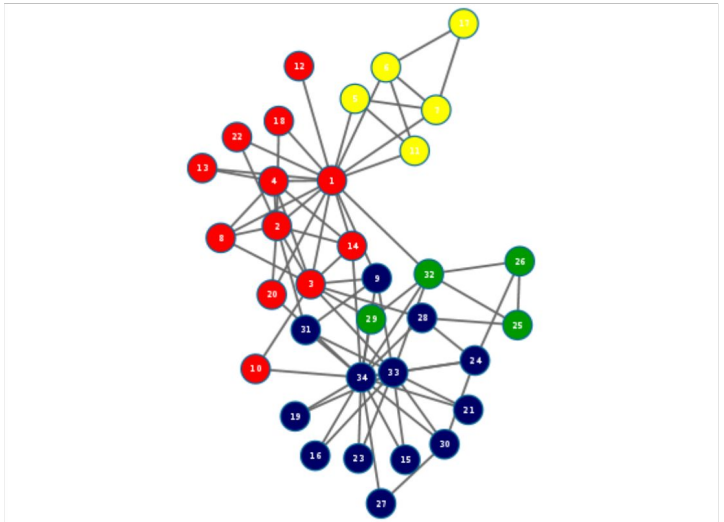
Four randomly generated with noise and overlapping nodes.

Karate Club

FLEX



Modularity

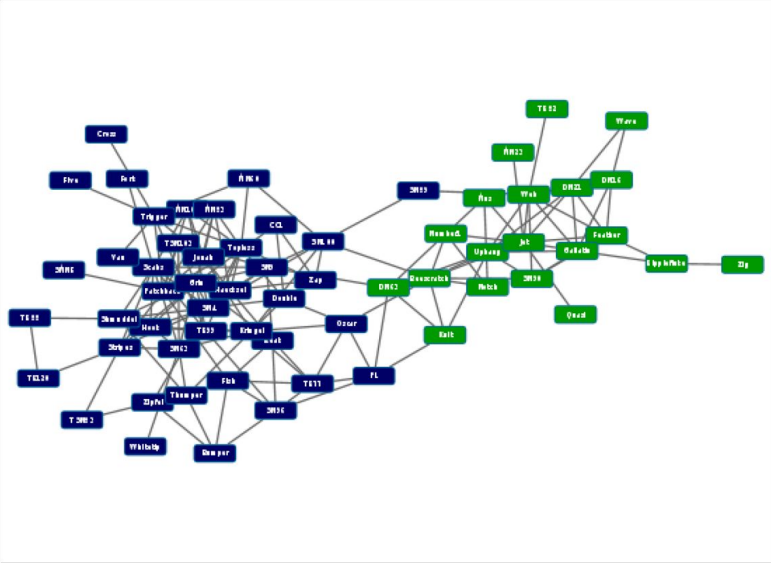


NMI **0.95** **0.40**

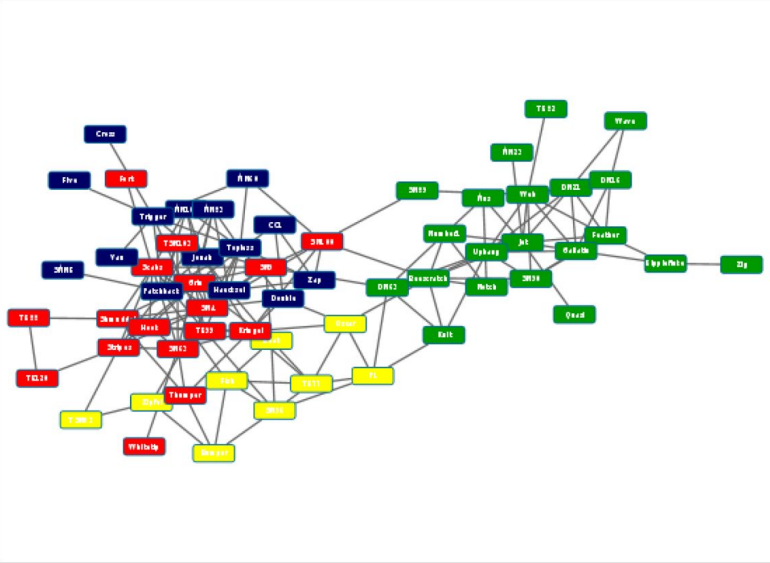
α=0.8, β=0.3, γ=2

Dolphins

FLEX



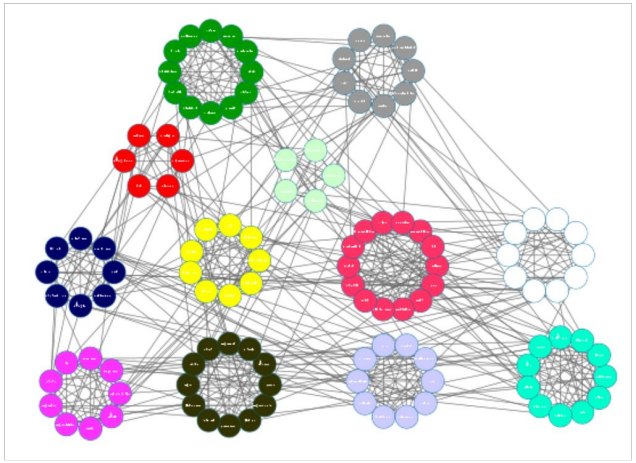
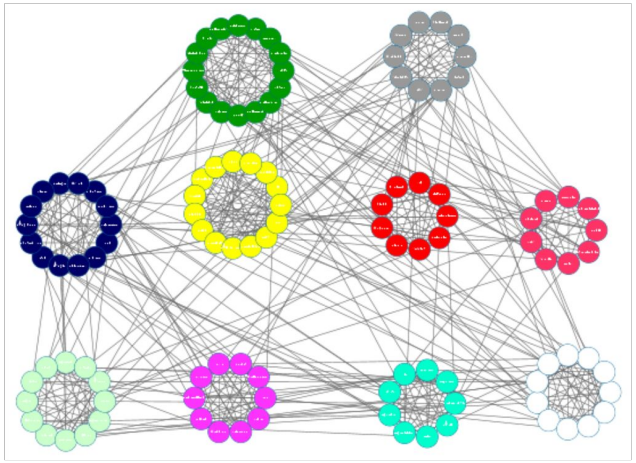
Modularity



NMI **0.86** **0.46**

α=0.4, β=0.3, γ=4

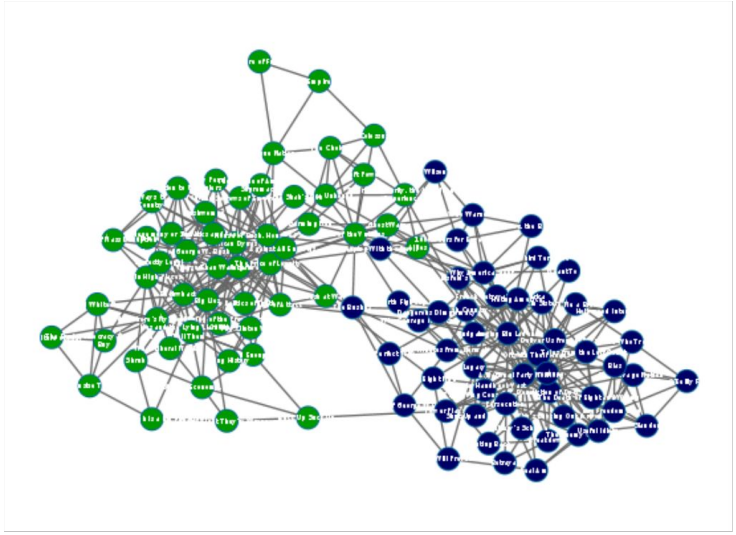
Football

	FLEX	Modularity
		
NMI	0.74	0.67

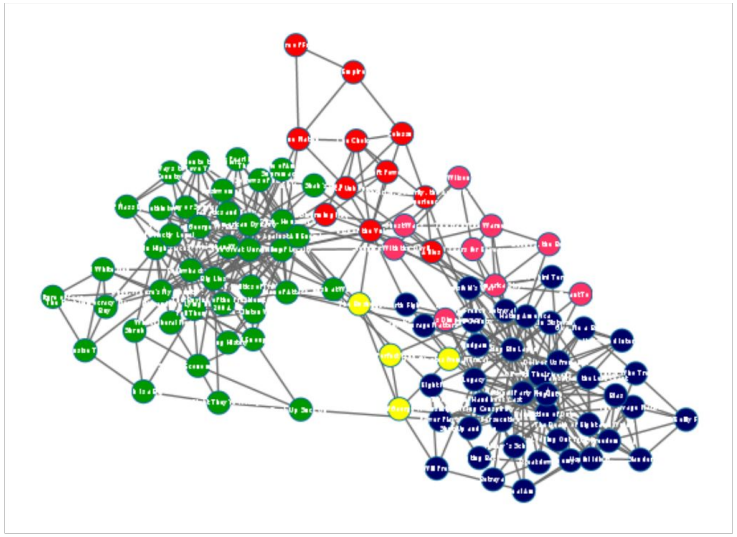
α=0.8, β=0.6, γ=4

Krebs

FLEX



Modularity



NMI **0.45** **0.32**

α=0.8, β=0.3, γ=4

Random Networks

Network 50

$(\alpha = 0.8, \beta = 0.3, \gamma = 2)$

	w/o Noise		Noise	
Metric	FLEX	Modularity	FLEX	Modularity
NMI	0.94	0.94	**0.77**	0.43
NMI Over.	1.00	1.00	**0.78**	0.43
NMI Multimodal	0.87	**0.93**	**0.69**	0.42

Network 100

$(\alpha = 0.8, \beta = 0.3, \gamma = 4)$

	w/o Noise		Noise	
Metric	FLEX	Modularity	FLEX	Modularity
NMI	0.95	0.95	**0.93**	0.58
NMI Over.	0.98	0.98	**0.94**	0.58
NMI Multimodal	**0.96**	0.95	**0.82**	0.43

Network 200

$(\alpha = 0.8, \beta = 0.3, \gamma = 4)$

	w/o Noise		Noise	
Metric	FLEX	Modularity	FLEX	Modularity
NMI	0.97	0.97	**0.93**	0.60
NMI Over.	1.00	1.00	**0.94**	0.62
NMI Multimodal	**0.98**	0.97	**0.81**	0.47

Network 500

$(\alpha = 0.8, \beta = 0.3, \gamma = 4)$

	w/o Noise		Noise	
Metric	FLEX	Modularity	FLEX	Modularity
NMI	0.99	0.99	**0.89**	0.55
NMI Over.	1.00	1.00	**0.91**	0.56
NMI Multimodal	0.99	0.99	**0.73**	0.40

Final Remarks

- FLEX is an improvement to Modularity.
- The three weighting parameters requires a prior study of the network.
- But they give some hints on the network dynamics.
- Insight to overlapping nodes.

Future Work

- Adapt Louvain Method to use FLEX
- Comparison with other overlapping detection heuristics
- Large-scale networks