8th International Symposium on Neural Networks – ISNN2011

Particle Competition and Cooperation for Uncovering Network Overlap Community Structure

Fabricio A. Breve[1] , Liang Zhao[1], Marcos G. Quiles[2], Witold Pedrycz[3,4], Jimming Liu[5]

[1] Department of Computer Science, Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, SP, Brazil
[2] Department of Science and Technology (DCT), Federal University of São Paulo (Unifesp), São José dos Campos, SP, Brazil
[3] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, T6R 2V4, Canada
[4] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
[5] Computer Science Department, Hong Kong Baptist University, Kowloon, Hong Kong

Outline

- Introduction
  - Community Detection
  - Overlap Nodes
- Proposed Method
  - Nodes and Particles Dynamics
  - Distance Tables
  - Random-Deterministic Walk
- Computer Simulations
  - Artificial Network
  - Real-World Network
- Conclusions

Community Detection

- Many networks are found to be divided naturally into communities or modules, therefore discovering of these communities structure became an important research topic.
- The problem of community detection is very hard and not yet satisfactorily solved, despite a large amount of efforts having been made over the past years.

[1] Newman, M.E.J., Girvan, M.: **Finding and evaluating community structure in networks**. *Physical Review E* 69, 026113 (2004)
[2] Newman, M.: **Modularity and community structure in networks**. *Proceedings of the National Academy of Science of the United States of America* 103, 8577–8582 (2006)
[3] Duch, J., Arenas, A.: **Community detection in complex networks using extremal optimization**. *Physical Review E* 72, 027104 (2005)
[4] Reichardt, J., Bornholdt, S.: **Detecting fuzzy community structures in complex networks with a potts model**. *Physical Review Letters* 93(21), 218701 (2004)
[5] Danon, L., D´ıaz-Guilera, A., Duch, J., Arenas, A.: **Comparing community structure identification**. *Journal of Statistical Mechanics: Theory and Experiment* 9, P09008 (2005)
[6] Fortunato, S. **Community detection in graphs**. *Physics Reports* 486(3-5), 75–174 (2010)

Overlap Nodes

- There are common cases where some nodes in a network can belong to more than one community
  - Example: In a social network of friendship, individuals often belong to several communities: their families, their colleagues, their classmates, etc
  - These are called *overlap nodes*
  - Most known community detection algorithms do not have a mechanism to detect them

[7] Zhang, S., Wang, R.S., Zhang, X.S.: **Identification of overlapping community structure in complex networks using fuzzy c-means clustering**. *Physica A Statistical Mechanics and its Applications* (2007)

[8] Palla, G., Derényi, I., Farkas, I., Vicsek, T.: **Uncovering the overlapping community structure of complex networks in nature and society**. *Nature* (7043), 814–818 (2005)

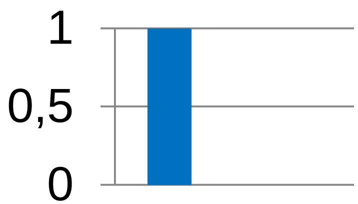
[9] Zhang, S., Wang, R.S., Zhang, X.S.: **Uncovering fuzzy community structure in complex networks**. *Physical Review E* 76(4), 046103 (2007)

Proposed Method

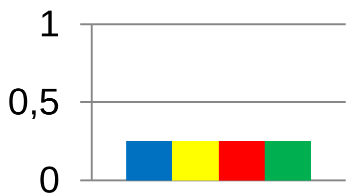
- Particles competition and cooperation in networks
  - Competition for possession of nodes of the network
  - Cooperation among particles from the same team (label)
    - Each team of particles tries to dominate as many nodes as possible in a cooperative way and at the same time prevent intrusion of particles of other teams.
  - Random-deterministic walk

Initial Configuration

- A particle is generated for each labeled node of the network
  - The node will be called that particle's *home node*
- Particles initial position are set to their respective home nodes.
- Particles with same label play for the same team
- Nodes have a domination vector
  - Labeled nodes have ownership set to their respective teams.
  - Unlabeled nodes have levels set equally for each team



Ex: [1 0 0 0]
(4 classes, node labeled as class A)

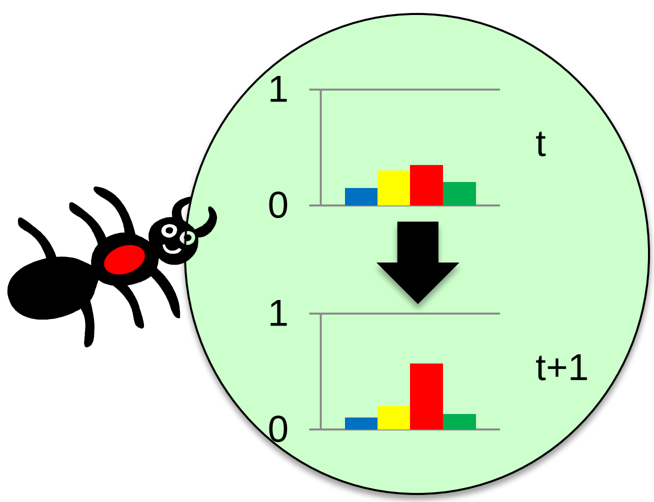


Ex: [0.25 0.25 0.25 0.25]
(4 classes, unlabeled node)

$$v_i^{\omega_\ell}(0) = \begin{cases} 1 & \text{if } y_i = \ell \\ 0 & \text{if } y_i \neq \ell \text{ and } y_i \in L \\ \frac{1}{c} & \text{if } y_i = \emptyset \end{cases}$$

Node Dynamics

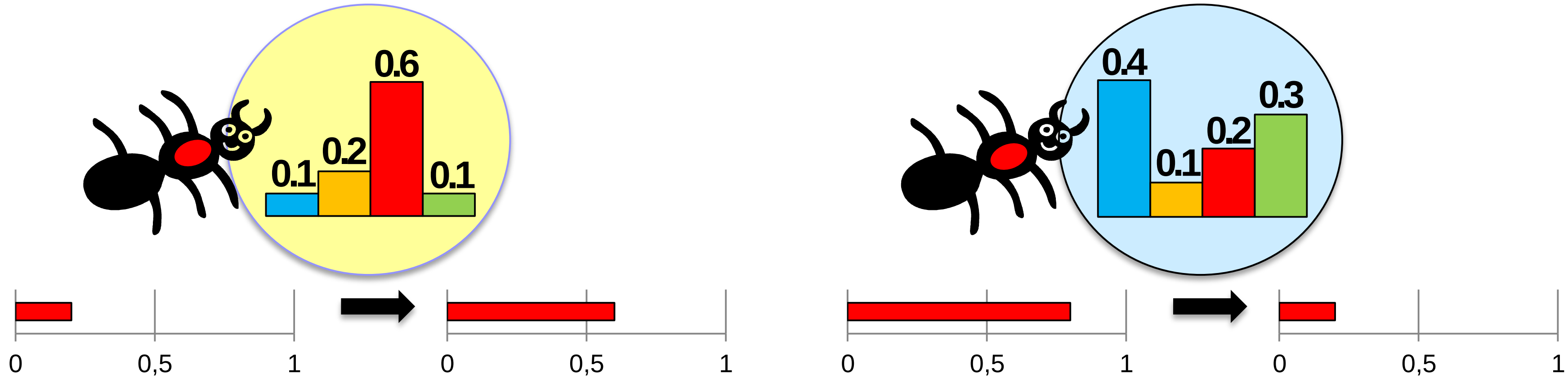
- When a particle selects a neighbor to visit:
  - It decreases the domination level of the other teams
  - It increases the domination level of its own team
  - Exception:
    - Labeled nodes domination levels are fixed



$$v_i^{\omega_\ell}(t+1) = \begin{cases} \max\{0, v_i^{\omega_\ell}(t) - \frac{\Delta_v \rho_j^{\omega}(t)}{c-1}\} \\ \quad \text{if } y_i = \emptyset \text{ and } \ell \neq \rho_j^f \\ v_i^{\omega_\ell}(t) + \sum_{q \neq \ell} v_i^{\omega_q}(t) - v_i^{\omega_q}(t+1) \\ \quad \text{if } y_i = \emptyset \text{ and } \ell = \rho_j^f \\ v_i^{\omega_\ell}(t) \quad \text{if } y_i \in L \end{cases}$$

Particle Dynamics

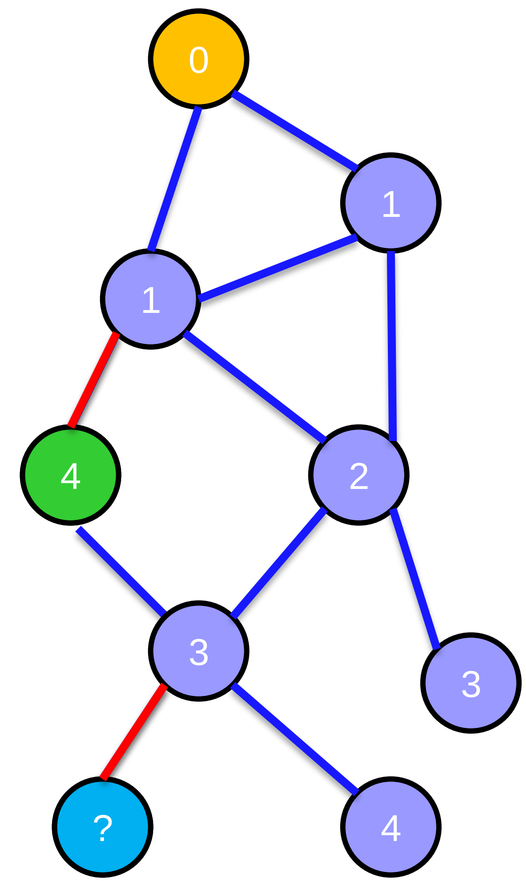
- A particle gets:
  - stronger when it selects a node being dominated by its team
  - weaker when it selects node dominated by other teams



$$\rho_j^{\omega}(t+1) = v_i^{\omega_\ell}(t+1).$$

Distance Table

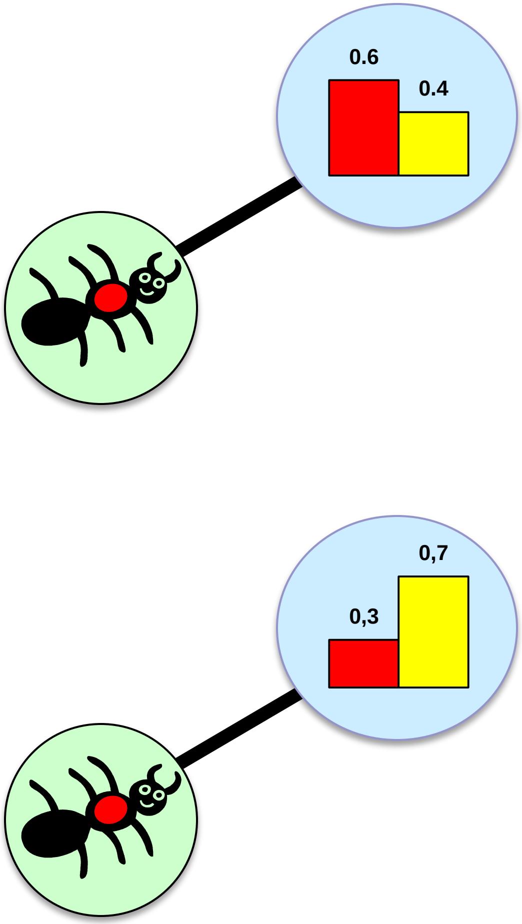
- Keep the particle aware of how far it is from its home node
  - Prevents the particle from losing all its strength when walking into enemies neighborhoods
  - Keep them around to protect their own neighborhood.
- Updated dynamically with local information
  - Does not require any prior calculation



$$\rho_j^{d_k}(t+1) = \begin{cases} \rho_j^{d_i}(t) + 1 \text{ if } \rho_j^{d_i}(t) + 1 < \rho_j^{d_k}(t) \\ \rho_j^{d_k}(t) \quad \text{otherwise} \end{cases}$$

Particles Walk

- Shocks
  - A particle really visits the selected node only if the domination level of its team is higher than others;
  - otherwise, a shock happens and the particle stays at the current node until next iteration.
- How a particle chooses a neighbor node to target?
  - Random walk
  - Deterministic walk



Random-Deterministic Walk

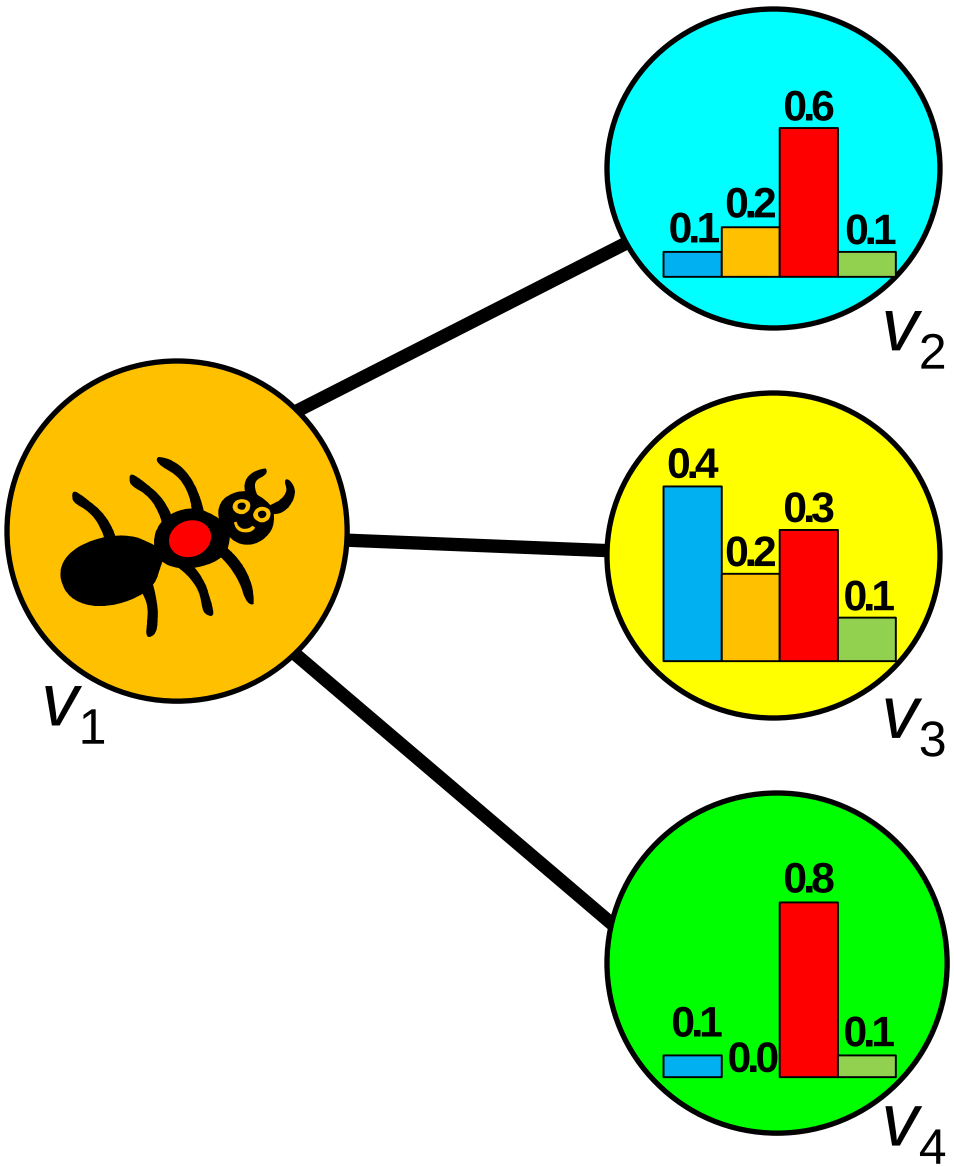
- Random walk
  - The particle randomly chooses any neighbor to visit with no concern about domination levels or distance

$$p(v_i|\rho_j) = \frac{W_{qi}}{\sum_{\mu=1}^{n} W_{q\mu}}$$

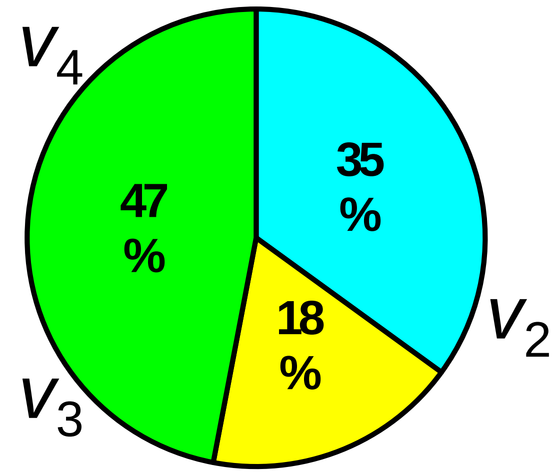
- Deterministic walk
  - The particle will prefer visiting nodes that its team already dominates and nodes that are closer to their home nodes

$$p(v_i|\rho_j) = \frac{W_{qi} v_i^{\omega_\ell} \frac{1}{(1+\rho_j^{d_i})^2}}{\sum_{\mu=1}^{n} W_{q\mu} v_i^{\omega_\ell} \frac{1}{(1+\rho_j^{d_i})^2}}$$

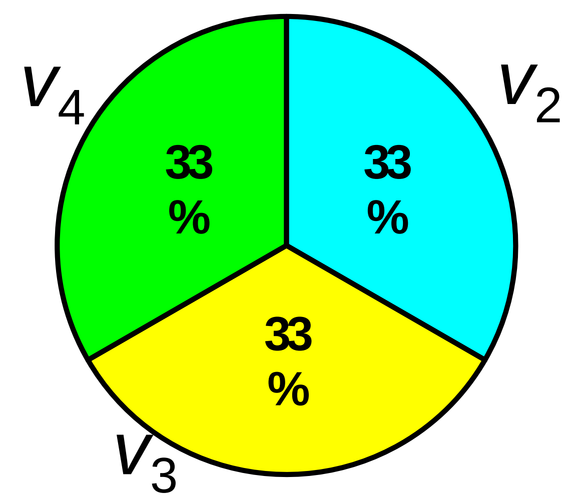
The particles must exhibit both movements in order to achieve an equilibrium between exploratory and defensive behavior



Deterministic Moving Probabilities

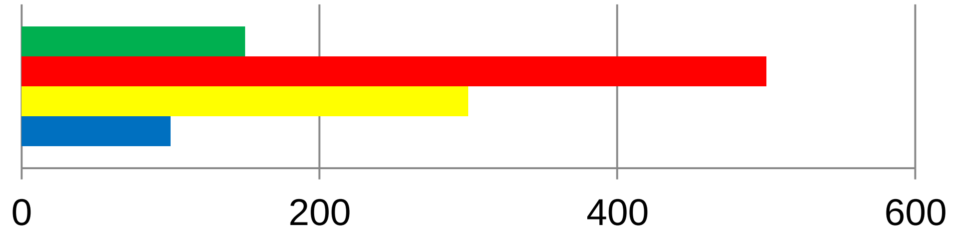


Random Moving Probabilities



Long Term Domination Levels

- Each time a particle visits a node using **random walk**, it also increases its team *long term domination levels* accordingly to its strength.
  - All levels starts from zero
  - No upper limit
  - No decrease in other team levels



$$v_i^{\lambda_\ell}(t+1) = v_i^{\lambda_\ell}(t) + \rho_j^{\omega}(t)$$

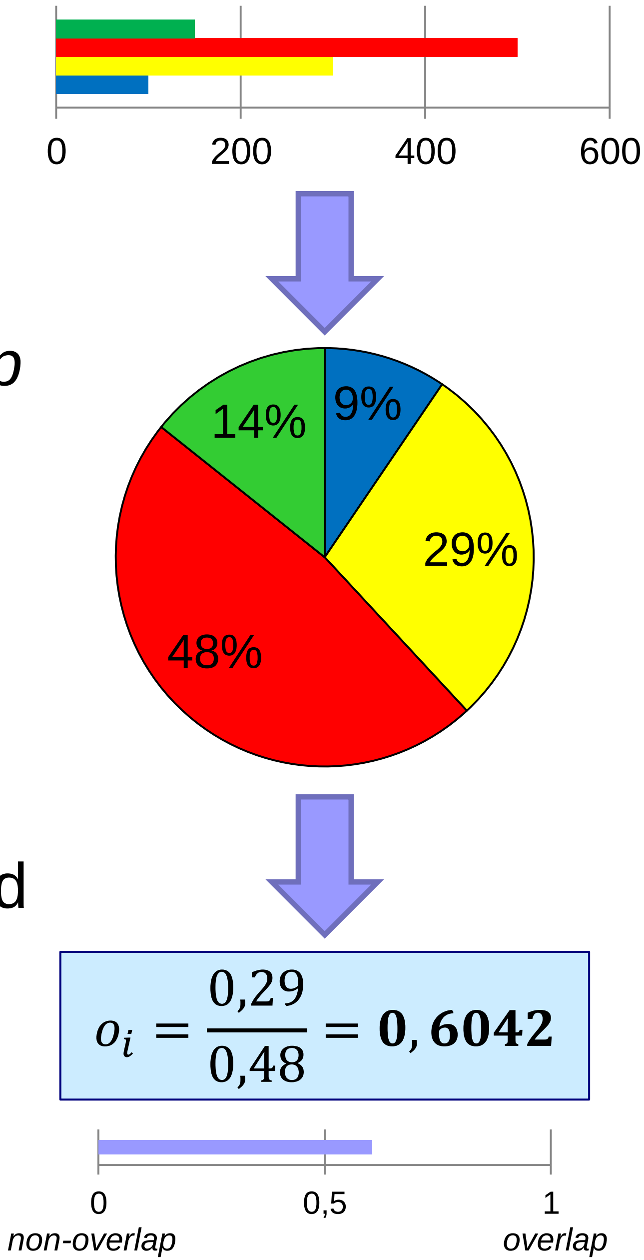
Fuzzy Output and Overlap Indexes

- After the last iteration, the *membership degrees* are calculated based on *long term domination levels*

$$f_i^{\ell} = \frac{v_i^{\lambda_\ell}(\infty)}{\sum_{q=1}^{c} v_i^{\lambda_q}(\infty)}$$

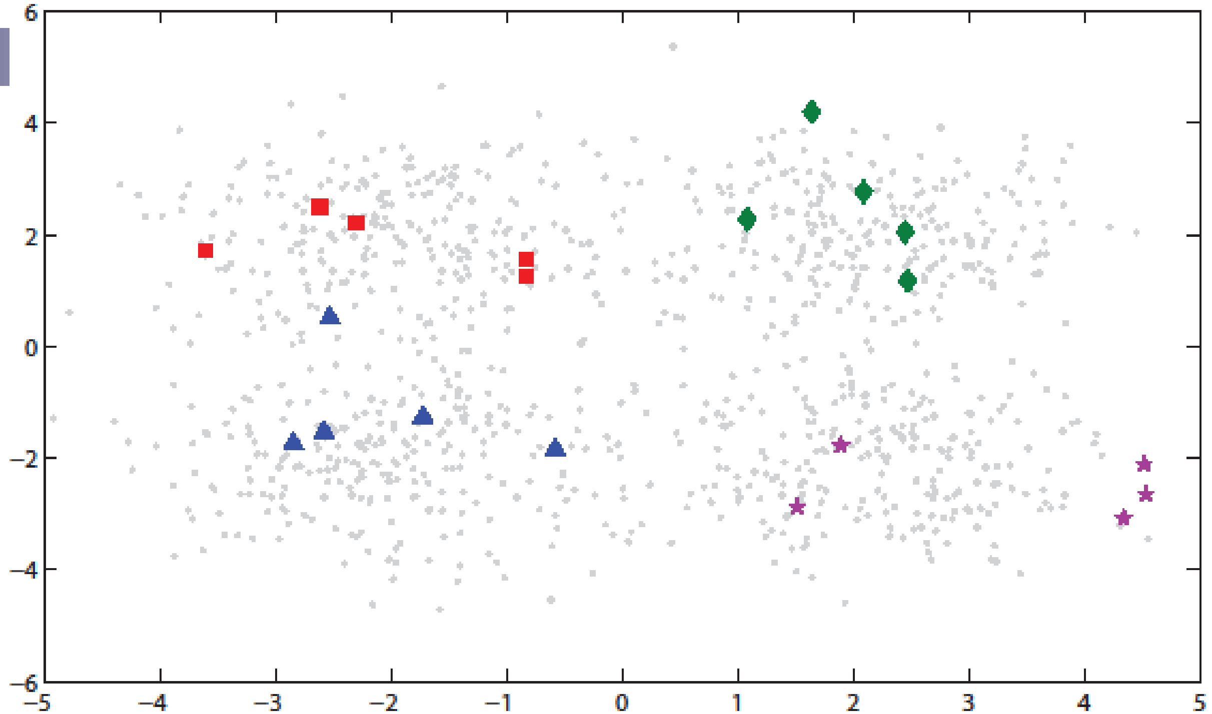
- And the *overlap indexes* are calculated from the *membership degrees*

$$o_i = \frac{f_i^{\ell**}}{f_i^{\ell*}}$$

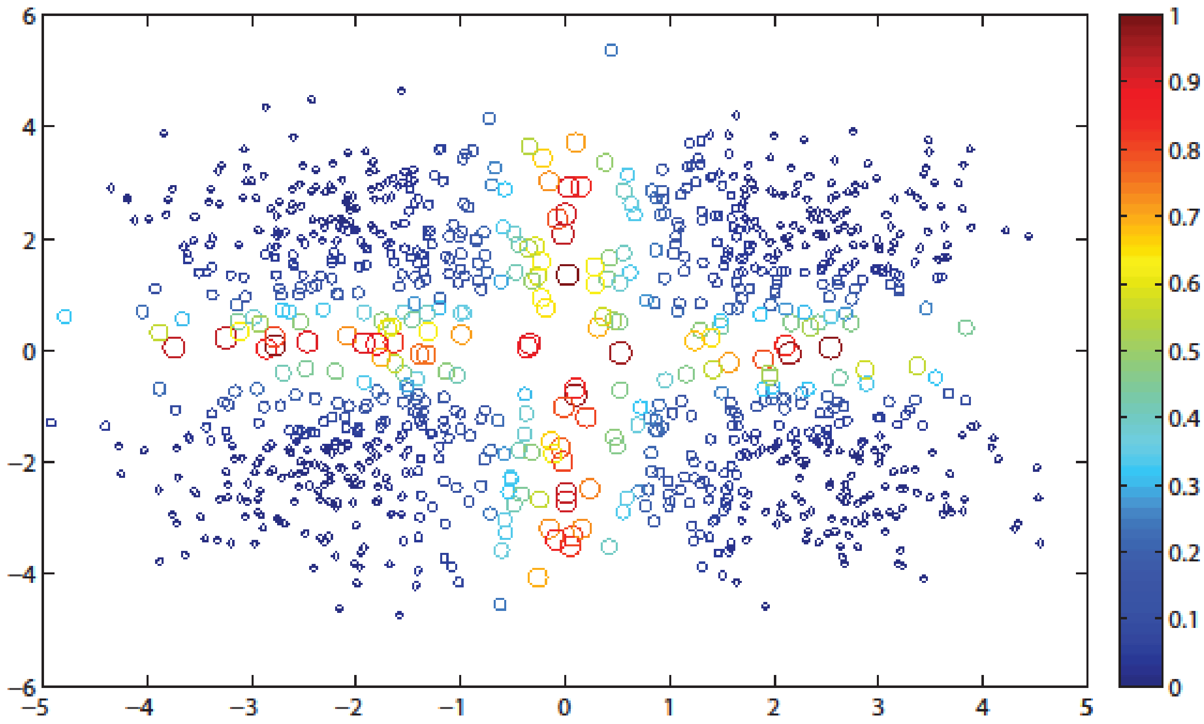


Computer simulations:

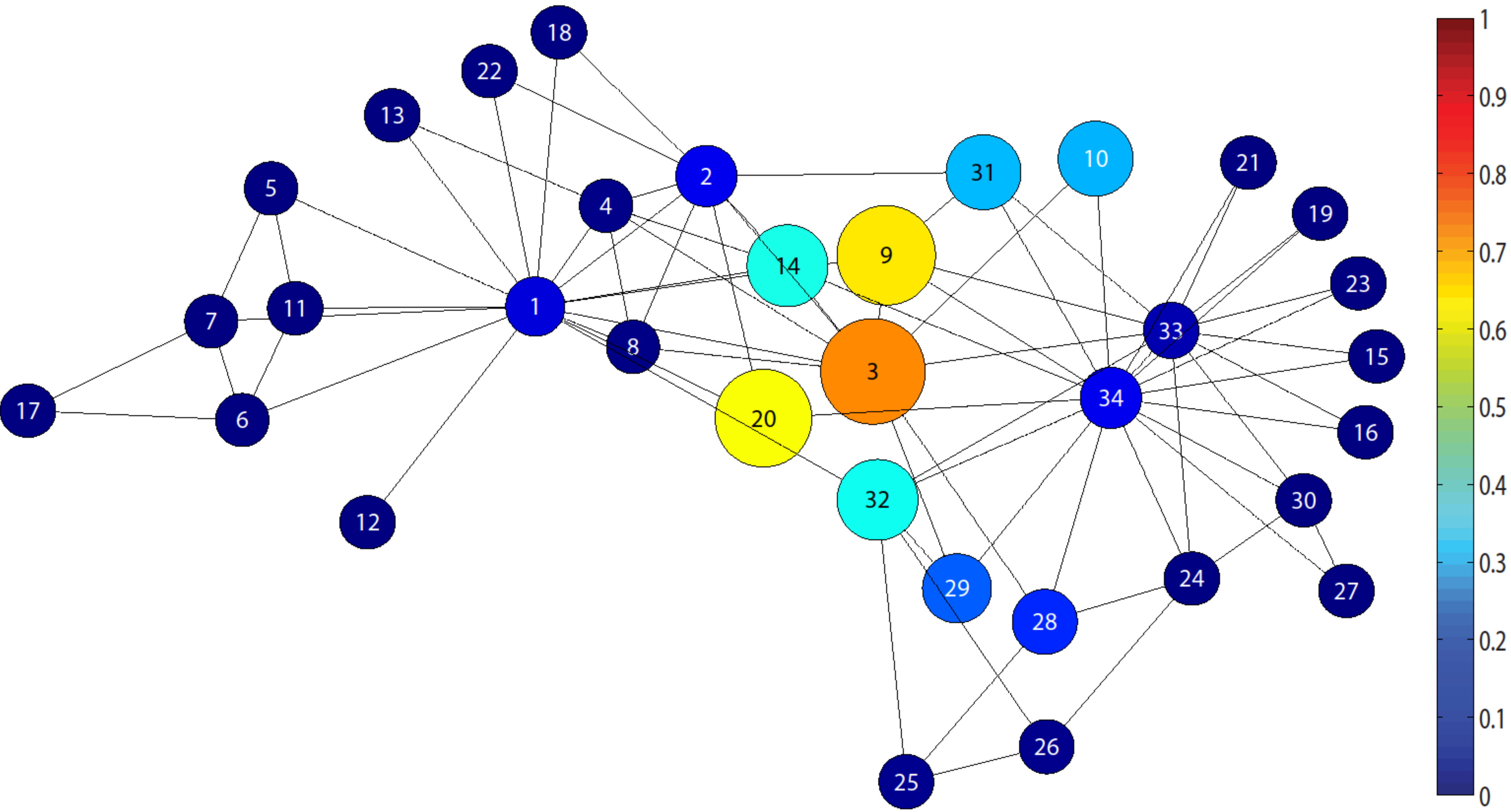
Classification of normally distributed classes (Gaussian distribution)



(a) toy data set with 1, 000 samples divided in four classes, 20 samples are labeled, 5 from each class (red squares, blue triangles, green lozenges and purple stars).



(b) nodes size and colors represent their respective overlap index detected by the proposed method.



Computer Simulations: The karate club network. Nodes size and colors represent their respective overlap index detected by the proposed method. Nodes 1 and 34 are pre-labeled.

Conclusions

- New semi-supervised learning graph-based method for uncovering the network overlap community structure.
  - It combines cooperation and competition among particles in order to generate a fuzzy output (soft label) for each node in the network
  - The fuzzy output correspond to the levels of membership of the nodes to each class
  - An overlap measure is derived from these fuzzy output, and it can be considered as a confidence level on the output label

Acknowledgements

- This work was supported by the State of São Paulo Research Foundation (FAPESP) and the Brazilian National Council of Technological and Scientific Development (CNPq)

8th International Symposium on Neural Networks – ISNN2011

Particle Competition and Cooperation for Uncovering Network Overlap Community Structure

Fabricio A. Breve[1] , Liang Zhao[1], Marcos G. Quiles[2], Witold Pedrycz[3,4], Jimming Liu[5]

[1] Department of Computer Science, Institute of Mathematics and Computer Science (ICMC), University of São Paulo (USP), São Carlos, SP, Brazil
[2] Department of Science and Technology (DCT), Federal University of São Paulo (Unifesp), São José dos Campos, SP, Brazil
[3] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, T6R 2V4, Canada
[4] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
[5] Computer Science Department, Hong Kong Baptist University, Kowloon, Hong Kong