Visualizing Communities in Dynamic Networks

M.G. Beiró, J.R. Busch, J.I. Alvarez-Hamelin

Facultad de Ingeniería UBA

November 4th, 2010

Complex Networks and Data Communications Group

Topics

Static community detection

Dynamic communities

Visualization

A social blogs network

Conclusions

Future work

Topics

Static community detection

Dynamic communities

Visualization

A social blogs network

Conclusions

Future work

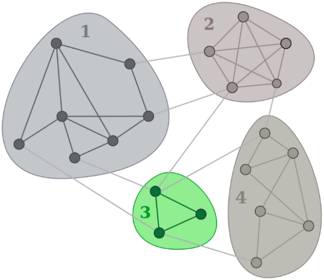
Modularity

- Modularity is a standard measure of quality of a community structure

$$k(C) = \sum_{v \in C} k(v) \tag{1}$$

$$n(C) = \sum_{v,w \in C} m(v,w) \tag{2}$$

$$Q(\mathcal{C}) = \sum_{C \in \mathcal{C}} \left(\frac{n(C)}{k(V)} - \frac{k^2(C)}{k^2(V)} \right) \tag{3}$$



Modularity optimization

(Brandes *et al.*, "On modularity clustering", 2008)

Determining if a given partition is optimal is NP-complete

Several algorithms

- Louvain (Blondel *et al.*, greedy)
- Newman-Girvan (greedy)
- Duch-Arenas (extremal optimization)

Fortunato & Barthelemy, "Resolution limit in community detection", 2006

Modularity has a resolution limit

Resolution-based algorithms

- Reichardt-Bornholdt (simmulated annealing)
- Busch *et. al* (submodularity, greedy)
- Aynaud (multi-scale, greedy)

Topics

Static community detection

Dynamic communities

Visualization

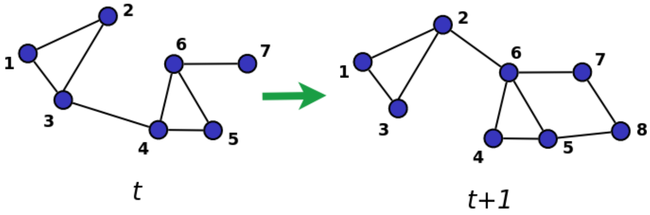
A social blogs network

Conclusions

Future work

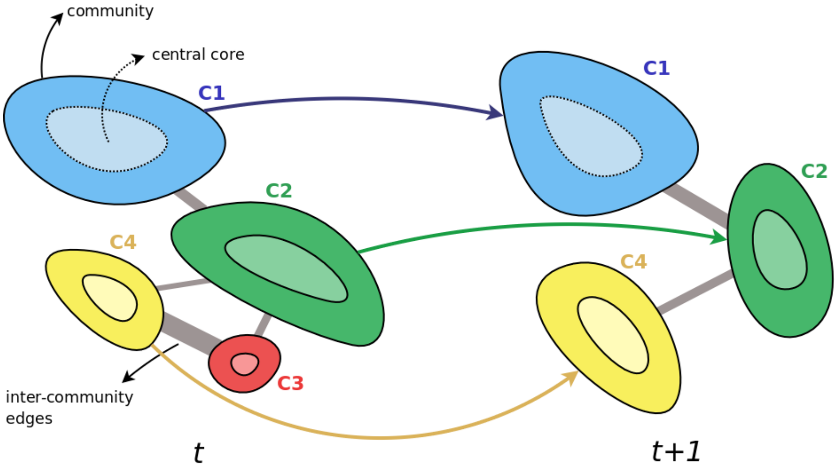
Dynamic networks

- The network evolves through time
- New nodes are added, some disappear
- Edges may change also
- Two approaches arise
  - A quality function including temporal information
  - Tracking of static communities through time



Dynamic communities

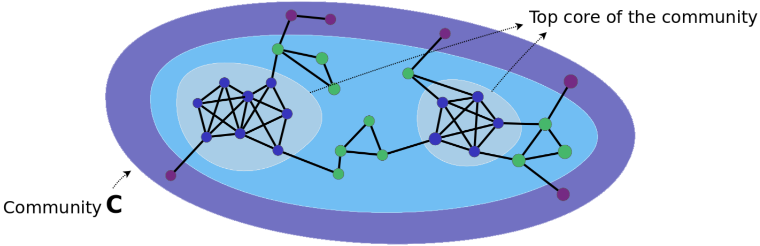
Our approach



Dynamic communities

Our approach

- Central hubs
  - Detected through *k-core decomposition*
- Similarity
  - Communities are matched one-to-one between t and $t+1$
  - $s(C_t, C_{t+1}) = |K(C_t) \cap K(C_{t+1})|$
  - The pairs with bigger similarity are joined first
  - Some communities in t may find no peer in $t+1$



Dynamic communities

Other approaches

- Hopcroft, "Tracking evolving communities in large linked networks"
  - First work in the area, 2004
  - Uses *cosine similarity* as a measure
- Palla *et al.*, "Quantifying social group evolution"
  - Finds communities J in G=G(t)+G(t+1)
  - Then it projects them to t and $t+1$
  - Picks the community in t with biggest *relative* overlap with J
  - Same with $t+1$
  - $o(C_t, C_{t+1}) = \frac{|(C_t) \cap (C_{t+1})|}{|(C_t) \cup (C_{t+1})|}$

Topics

Static community detection

Dynamic communities

Visualization

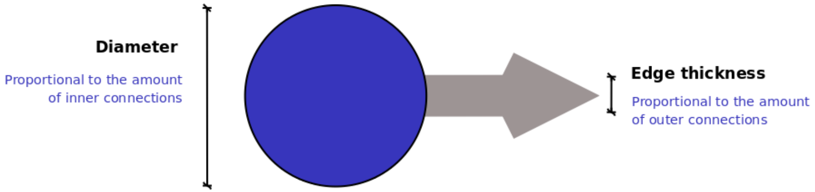
A social blogs network

Conclusions

Future work

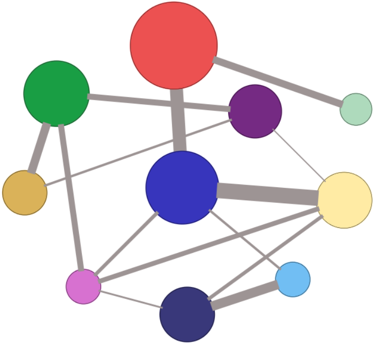
SnailVis

- SnailVis is a software to visualize communities
- Scales to large networks
- Provides an abstraction of the graph, based on the partition into communities



SnailVis

- Communities are rendered following a spiral
- They keep a minimum distance and do not overlap



$$\rho = K \cdot \theta^{\beta}, \beta \in R \quad (4)$$

(*Fermat*'s spiral)

SnailVis for Dynamics

- We adapted SnailVis to visualize dynamic networks
- Foresee the maximum size each community will reach
- Give to each community the space it will need
- Communities remain fixed through time

Topics

Static community detection

Dynamic communities

Visualization

A social blogs network

Conclusions

Future work

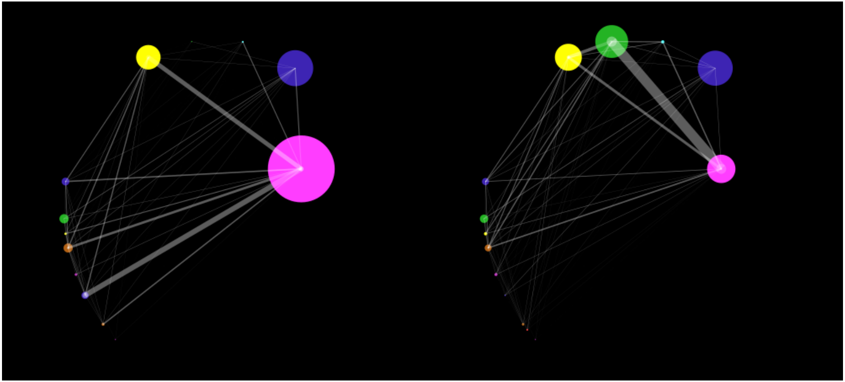
Case Study: A dynamic blogs dataset

- 3,772 nodes
- 36,750 edges

- Each node represents a **blog**
- Blogs are connected by **links** (between their articles)
- Links are persistent
- Exploration done on a **daily** basis, during **four months**
- **120 snapshots**

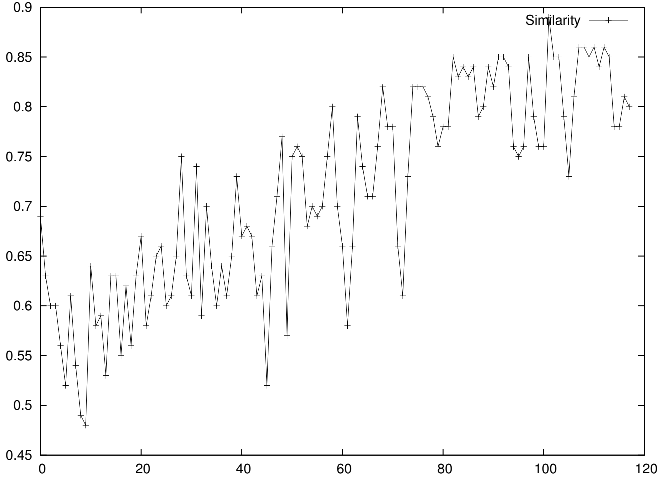
Data obtained from the ANR WebFluence project

Results



Results

Similarity through time



Topics

Static community detection

Dynamic communities

Visualization

A social blogs network

Conclusions

Future work

Conclusions

- We find static communities with the submodularity algorithm
- Static communities are matched one-to-one to track each community's life
- Identified communities are visualized in a spiral
  - They keep a fixed position through their life
  - They may grow or get smaller
  - They do not overlap in space
- We analyze a dynamic blogs network.

Topics

Static community detection

Dynamic communities

Visualization

A social blogs network

Conclusions

Future work

Future work

- Consider other possible processes during a community life (Palla *et al.*, 2007)
  - Merging
  - Spliting
- Analyze stability and compare with other methods