Background & Objectives

- **Folksonomy**: Users annotate online Resources with Tags
- **Folksonomy is modeled as Tripartite Hypergraph**
- Almost all existing community detection algorithms for folksonomies assign only a single community to each node
- **Reality**: Nodes belong to multiple overlapping communities
  - Most users have multiple topics of interest
  - The same resource is often associated with semantically different tags by different users
- Two prior approaches exist for overlapping community detection. Both work on projection of the tripartite hypergraph.
- Projections lose information and quality of communities is proved to be worse in projected network.
- **Our Objective**: Develop an algorithm to detect overlapping communities in folksonomies considering the complete tripartite hypergraph structure

Algorithm

**Idea:** Cluster links in stead of nodes

- Find similarities between all pairs of adjacent hyperedges using Algorithm 1.
- Construct the weighted line graph of the hypergraph. Hyperedges are nodes here and two such nodes are connected if they have non-zero similarity. Exact similarity score is represented as the edge-weight.
- Apply any community detection algorithm on that line graph. (We used Infomap algorithm [Rosvall et al., PNAS 2008])
- Each hyperedge gets placed into a single link-community.
- A node inherits membership of all those communities into which the hyperedges connected with this node are placed.

**Time Complexity = O(n.a.d^2)** where n = number of nodes and d = average degree

Motivation

**Why Overlapping Communities?**

- Existing algorithms likely to put the daffodil image only into ‘Daffodil’ community based on majority tagging
- Algorithm for overlapping community detection relate image with ‘Yellow’ community as well, can be recommended to users favoring yellow objects ⇒ better community-based recommendation
- Identify ‘Daffodil’ community as a subset of ‘Flower’ community ⇒ hierarchical organization of resources and tags into semantic categories

Experiments

- Synthetic hypergraphs generated
  - Each node assigned to one community, then β fraction of nodes assigned multiple communities
  - Nodes in same community randomly connected with hyperedges
  - Number of hyperedges is decided based on the specified density α

Conclusion

- We proposed the first algorithm to detect overlapping communities considering the full tripartite hypergraph structure of folksonomies.
- It outperforms existing algorithms that consider projections of hypergraphs.
- The proposed algorithm can be used in recommending interesting resources and friends to users.

Contact

Abhijnan Chakraborty: abhijnan.cs@gmail.com
Saptarshi Ghosh: saptarshi.ghosh@gmail.com
Niloy Ganguly: niloy@cs.iitkgp.ernet.in

Complex Network Research Group (CNeRG)
Department of Computer Science & Engineering
Indian Institute of Technology Kharagpur
http://cse.iitkgp.ac.in/resgrp/cnerg