
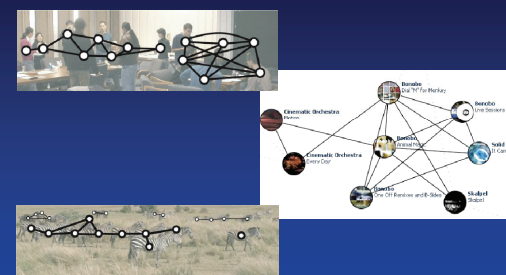


A Framework for Finding Communities in Dynamic Social Networks

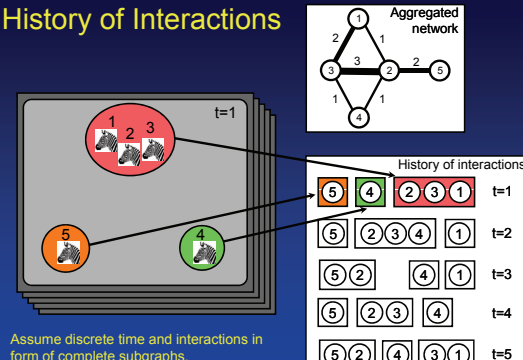


Slides from Chayant Tantipathananandh, Tanya Berger-Wolf
University of Illinois at Chicago
David Kempe
University of Southern California

Social Networks



History of Interactions



Assume discrete time and interactions in form of complete subgraphs.

t=1	5	4	2	3	1
t=2	5	2	3	4	1
t=3	5	2	4	1	
t=4	5	2	3	4	
t=5	5	2	4	3	1

Community Identification

What is community?
 "Cohesive subgroups are subsets of actors among whom there are relatively strong, direct, intense, frequent, or positive ties." [Wasserman & Faust '97]

Notions of communities:

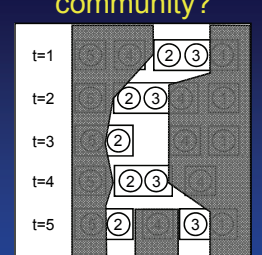
Static

- Centrality and betweenness [Girvan & Newman '01]
- Correlation clustering [Basal et al. '02]
- Overlapping cliques [Palla et al. '05]

Dynamic

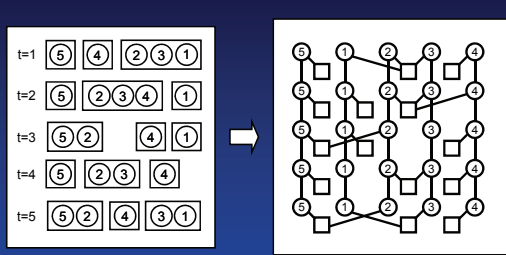
- Metagroups [Berger-Wolf & Saia '06]

The Question: What is dynamic community?



- A dynamic community is a subset of individuals that stick together over time.
- NOTE: Communities ≠ Groups

Approach: Graph Model



t=1	5	4	2	3	1
t=2	5	2	3	4	1
t=3	5	2	4	1	
t=4	5	2	3	4	
t=5	5	2	4	3	1

Approach: Assumptions

Required

- Individuals and groups represent exactly **one** community at a time.
- Concurrent groups represent **distinct** communities.

Desired

- **Conservatism:** community affiliation changes are rare.
- **Group Loyalty:** individuals observed in a group belong to the same community.
- **Parsimony:** few affiliations overall for each individual.

Approach: Color = Community

T1

T2

T3

T4

T5

Approach: Assumptions

Required

- Individuals and groups represent exactly **one** community at a time.
- Concurrent groups represent **distinct** communities.

Desired

- **Conservatism:** community affiliation changes are rare.
- **Group Loyalty:** individuals observed in a group belong to the same community.
- **Parsimony:** few affiliations overall for each individual.

Costs

- **Conservatism:** switching cost (α)
- **Group loyalty:**
 - Being absent (β_1)
 - Being different (β_2)
- **Parsimony:** number of colors (γ)

Approach: Assumptions

Required

- Individuals and groups represent exactly **one** community at a time.
- Concurrent groups represent **distinct** communities.

Desired

- **Conservatism:** community affiliation changes are rare.
- **Group Loyalty:** individuals observed in a group belong to the same community.
- **Parsimony:** few affiliations overall for each individual.

Costs

- **Conservatism:** switching cost (α)
- **Group loyalty:**
 - Being absent (β_1)
 - Being different (β_2)
- **Parsimony:** number of colors (γ)

Approach: Assumptions

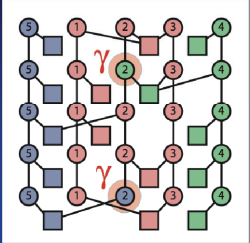
Required

- Individuals and groups represent exactly **one** community at a time.
- Concurrent groups represent **distinct** communities.

Desired

- Conservatism:** community affiliation changes are rare.
- Group Loyalty:** individuals observed in a group belong to the same community.
- Parsimony:** few affiliations overall for each individual.

Costs



- Conservatism: switching cost (α)
- Group loyalty:
 - Being absent (β_1)
 - Being different (β_2)
- Parsimony: number of colors (γ)

Problem Definition

- Minimum Community Interpretation**
For a given cost setting, $(\alpha, \beta_1, \beta_2, \gamma)$, find vertex coloring that minimizes total cost.
 - Color of group vertices = Community structure
 - Color of individual vertices = Affiliation sequences
- Problem is NP-Complete and APX-Hard
 - Reduction from Minimum Multiway Cut problem


Model Validation and Algorithms

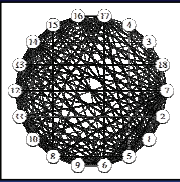
- Model validation: exhaustive search for an exact minimum-cost coloring
 - Fix group coloring and find optimal coloring for individuals using DP
- Heuristic algorithms evaluation: compare heuristic results to OPT.
 - Bipartite Matching Heuristic
 - Greedy Heuristics

Results

- Synthetic Data sets
 - Assembly Line
 - Dutiful Children
- Real-World Data sets
 - Southern Women
 - Grevy's Zebra Data set

Southern Women Data Set

by  Photograph by Ben Shain, Natchez, MS, October, 1935

 Aggregated network

Name of Participants in Game 1	One Player and Doves or Rabbits, Brown Beavers in DM City Run!																	
	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17
1. Mrs. Evelyn Johnson	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
2. Miss Anne MacFarlane	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
3. Miss Thelma Andrews	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
4. Miss Emma Rogers	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
5. Miss Charlotte McEwen	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
6. Miss Frances Anderson	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
7. Miss Gertrude Pitt	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
8. Miss Pearl Ogilvie	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
9. Miss Sarah McCall	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
10. Miss Vera Buchanan	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
11. Miss Vera Lillard	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
12. Miss Elizabeth Rogers	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
13. Miss Stella Arnold	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
14. Miss Don Taylor	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
15. Miss Helen Light	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
16. Miss Dorothy Almond	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
17. Miss Olive Colburn	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
18. Miss Vera Pitt	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Event participation

Ethnography

by Davis, Gardner, and Gardner, 1941

Name of Participant or Group	Cost Functions and Dates of Social Network Sessions by Old City Herald													
	4/10	5/1	6/10	6/20	6/29	7/10	7/20	7/29	8/10	8/20	8/29	9/10	9/20	9/29
1. Mrs. Evelyn Jefferson														
2. Miss Laura Mauldinville														
3. Miss Emma Anderson														
4. Miss Emma Rogers														
5. Miss Charlotte McIlwain														
6. Miss Frances Anderson														
7. Miss Eleanor Rye														
8. Miss Fred Ophthorp														
9. Miss Ruth DeSquad			X	X	X	X	X	X	X	X	X	X	X	X
10. Miss Vera Swadlow														
11. Miss Myra Liddell														
12. Miss Katherine Rogers														
13. Miss Sylvia Aronson														
14. Mrs. Nora Fayette														
15. Mrs. Helen Lloyd														
16. Mrs. Dorothy Marchison			X	X	X	X	X	X	X	X	X	X	X	X
17. Mrs. Clara Caldwell														
18. Mrs. Flora Price														

An Optimal Coloring:

$(\alpha, \beta, \gamma) = (1, 1, 3, 1)$

cost $(1, 1, 3, 1) * (25, 18, 4, 15) = 70$ 4 i-colors 3 g-colors

cost $(1, 1, 3, 1) * (25, 18, 4, 15) = 70$

Node	1	2	3	4	1	2	3	4	1	2	3	4
1	1	1	1	1	0	0	0	0	0	0	0	0
2	1	1	1	1	0	0	0	0	0	0	0	0
3	1	1	1	1	0	0	0	0	0	0	0	0
4	1	1	1	1	0	0	0	0	0	0	0	0
5	1	1	1	1	0	0	0	0	0	0	0	0
6	1	1	1	1	0	0	0	0	0	0	0	0
7	1	1	1	1	0	0	0	0	0	0	0	0
8	1	1	1	1	0	0	0	0	0	0	0	0
9	1	1	1	1	0	0	0	0	0	0	0	0
10	1	1	1	1	0	0	0	0	0	0	0	0
11	1	1	1	1	0	0	0	0	0	0	0	0
12	1	1	1	1	0	0	0	0	0	0	0	0
13	1	1	1	1	0	0	0	0	0	0	0	0
14	1	1	1	1	0	0	0	0	0	0	0	0
15	1	1	1	1	0	0	0	0	0	0	0	0
16	1	1	1	1	0	0	0	0	0	0	0	0
17	1	1	1	1	0	0	0	0	0	0	0	0
18	1	1	1	1	0	0	0	0	0	0	0	0

An Optimal Coloring:

$(\alpha, \beta, \gamma) = (1, 1, 1, 1)$

cost $(1, 1, 1, 1) * (2, 8, 29, 2) = 41$ 4 i-colors 3 g-colors

cost $(1, 1, 1, 1) * (2, 8, 29, 2) = 41$

Node	1	2	3	4	1	2	3	4	1	2	3	4
1	1	1	1	1	0	0	0	0	0	0	0	0
2	1	1	1	1	0	0	0	0	0	0	0	0
3	1	1	1	1	0	0	0	0	0	0	0	0
4	1	1	1	1	0	0	0	0	0	0	0	0
5	1	1	1	1	0	0	0	0	0	0	0	0
6	1	1	1	1	0	0	0	0	0	0	0	0
7	1	1	1	1	0	0	0	0	0	0	0	0
8	1	1	1	1	0	0	0	0	0	0	0	0
9	1	1	1	1	0	0	0	0	0	0	0	0
10	1	1	1	1	0	0	0	0	0	0	0	0
11	1	1	1	1	0	0	0	0	0	0	0	0
12	1	1	1	1	0	0	0	0	0	0	0	0
13	1	1	1	1	0	0	0	0	0	0	0	0
14	1	1	1	1	0	0	0	0	0	0	0	0
15	1	1	1	1	0	0	0	0	0	0	0	0
16	1	1	1	1	0	0	0	0	0	0	0	0
17	1	1	1	1	0	0	0	0	0	0	0	0
18	1	1	1	1	0	0	0	0	0	0	0	0

Conclusions

- An optimization-based framework for finding communities in dynamic social networks.
- Finding an optimal solution is NP-Complete and APX-Hard.
- Model evaluation by exhaustive search.
- Heuristic algorithms for larger data sets. Heuristic results comparable to optimal.

Discussion

- Scalability – datasets considered are small. What about large datasets? E.g., Facebook
- Assumption are too tight: Relaxation of their constraints
 - Individuals and groups represent exactly **one** community at a time
 - How does it effect scalability?
- What next after finding communities? Is their work motivating?

Thank You