Community Analysis

CS 579 Online Social Network Analysis

Dr. Cindy Hood
10/21/25

Homework Assignments

- HW #4 – Chicago Community Areas + Census Data
  - You may work in groups up to 4 students (no exceptions) on this hw
  - Due date now 11/3 <-- Change

- Please contact TAs with questions

HW #4 Using census data to explore Chicago Community Areas (CAs)

Each student has been assigned a Community Area (1-77) to explore. You will use data from census.gov to do an analysis of your CA using census data from the ACS (block group level) and decennial census (block level). Depending on the number of block groups in your CA, you may consider the broader geographic neighborhood of your CA as well (see instructions below). The deliverables are:

1. A data-based characterization of your CA including the recent history (back to 2010) of your CA

I suggest you use compare a few variables from the decennial census' in 2010 and 2020+ decennial or later ACS survey

- Population
- Demographics

2. A description of the data (i.e. variables) that you have chosen to use as the basis of your study along with a discussion of why you chose these variables and any preliminary analysis you did to narrow the data. You will choose at least 6 variables that are available at the block group level (i.e. from the ACS).
3. An analysis of similarity of the block groups that comprise your CA. If your CA has less than 60 block groups, you will add the block groups of geographic neighbors to get at least 60 block groups. This analysis should include modeling the block groups as a network.
    a. You will describe the modeling and analysis you did, stating assumptions and justifying decisions.
    b. You will provide data-based arguments including visualizations to support why your CA is a community.
    c. You will provide data-based arguments including visualizations to support why your CA is not a community.
    d. You will provide a proposal for alternative community(ies). If your CA has 60 or more block groups, this will be a proposal for organization of communities within your CA. If your CA has <60 block groups, you will propose an organization of block groups from your CA and neighbors (resulting in analysis of 60 or more). This proposal should include visualizations.
4. If you are working in a team, there is an additional step where you will put all your data together and propose an organization of communities. Each team member must have a unique CA so if your team has duplicates, please email me asap.
5. The above should be compiled into a report that includes citations and transcripts of any AI assistance. You will submit a pdf and code.

Teams will submit one report. Be sure to describe which Cas each team member analyzed

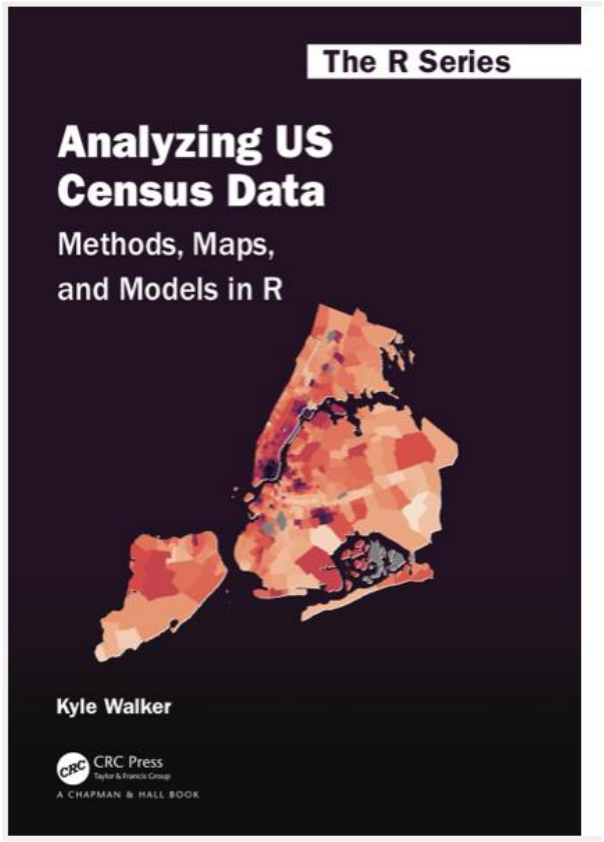
Exams and Final Project Poster Presentation

- Exam 1 – Grading underway
  - Target is to post grades by 10/27
  - Hand back exams in class on 10/28
- Exam 2 – Dec 2 in class
- Final Project Poster Session – Dec 4 in class
- Online students (sections 2 and 3) will have remote options

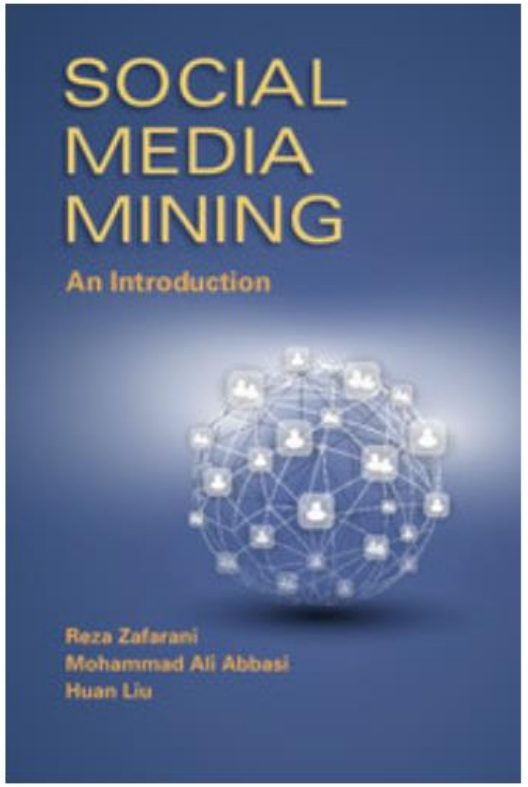
Final Project

- Extension of HW 4
- Teams of up to 4 permitted
- Proposal due 11/4

References



https://walker-data.com/census-r/



http://www.socialmediamining.info

Some additional resources

- Myatt and Johnson (2014), *Making Sense of Data I*, 2nd Edition, Wiley, ISBN: 978-1-118-40741-7
  - https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781118422007
- Networks, Crowds, and Markets: Reasoning About a Highly Connected World by David Easley and Jon Kleinberg.
  - http://www.cs.cornell.edu/home/kleinber/networks-book/
- Hanneman, Robert A. and Mark Riddle. 2005. Introduction to social network methods. Riverside, CA: University of California, Riverside
  - https://faculty.ucr.edu/~hanneman/nettext/

Recall ACS 2023 data at the block group level Median Household Income for all the Block Groups in Cook County

	GEOID	NAME	variable	estimate	moe
1	170310101001	Block Group 1; Census Tract 101; Cook County; Illinois	B29004_001	80809	36333
2	170310101002	Block Group 2; Census Tract 101; Cook County; Illinois	B29004_001	26351	9125
3	170310101003	Block Group 3; Census Tract 101; Cook County; Illinois	B29004_001	89545	33643
4	170310102011	Block Group 1; Census Tract 102.01; Cook County; Illi...	B29004_001	78700	23637
5	170310102012	Block Group 2; Census Tract 102.01; Cook County; Illi...	B29004_001	42176	8400
6	170310102013	Block Group 3; Census Tract 102.01; Cook County; Illi...	B29004_001	47569	15883
7	170310102014	Block Group 4; Census Tract 102.01; Cook County; Illi...	B29004_001	*NA*	*NA*
8	170310102021	Block Group 1; Census Tract 102.02; Cook County; Illi...	B29004_001	45590	10800
9	170310102022	Block Group 2; Census Tract 102.02; Cook County; Illi...	B29004_001	86419	16794
10	170310103001	Block Group 1; Census Tract 103; Cook County; Illinois	B29004_001	54130	34420
11	170310103002	Block Group 2; Census Tract 103; Cook County; Illinois	B29004_001	74688	27405
12	170310103003	Block Group 3; Census Tract 103; Cook County; Illinois	B29004_001	44335	20324
13	170310103004	Block Group 4; Census Tract 103; Cook County; Illinois	B29004_001	74500	17099
14	170310104001	Block Group 1; Census Tract 104; Cook County; Illinois	B29004_001	63693	27897

Showing 1 to 15 of 4,002 entries, 5 total columns

R code

```
med_house_income <- get_acs(
  geography = "block group",
  variables = "B29004_001",
  state = "IL",
  county = "Cook",
  year = 2023
)
```

```
Getting data from the 2019-2023 5-year ACS
Using FIPS code '17' for state 'IL'
Using FIPS code '031' for 'Cook County'
```

B29002_001 is the code for Median Household Income

Now I can join this table with the ACS Household Income Data

```
hhi_cca <- left_join(allchi_ca_bgs, med_house_income, by="GEOID")
```

NAME	variable	estimate	moe	geometry
Block Group 2; Census Tract 4005; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.62556 4*
Block Group 6; Census Tract 4804; Cook County; Illin...	B29004_001	100365	49442	*MULTIPOLYGON (((-87.58553 4*
Block Group 7; Census Tract 4804; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.58557 4*
Block Group 4; Census Tract 608; Cook County; Illinois	B29004_001	73356	7025	*MULTIPOLYGON (((-87.64794 4*
Block Group 1; Census Tract 2909; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.73006 4*
Block Group 2; Census Tract 2909; Cook County; Illin...	B29004_001	60072	28786	*MULTIPOLYGON (((-87.72995 4*
Block Group 3; Census Tract 2909; Cook County; Illin...	B29004_001	60833	35237	*MULTIPOLYGON (((-87.73729 4*
Block Group 4; Census Tract 2909; Cook County; Illin...	B29004_001	43490	22556	*MULTIPOLYGON (((-87.73993 4*
Block Group 2; Census Tract 8387; Cook County; Illin...	B29004_001	45139	29580	*MULTIPOLYGON (((-87.72506 4*
Block Group 1; Census Tract 3102; Cook County; Illin...	B29004_001	120263	22156	*MULTIPOLYGON (((-87.64539 4*
Block Group 2; Census Tract 3102; Cook County; Illin...	B29004_001	131250	39622	*MULTIPOLYGON (((-87.64671 4*
Block Group 2; Census Tract 3201.01; Cook County; I...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.62086 4*
Block Group 1; Census Tract 3201.01; Cook County; I...	B29004_001	158651	64476	*MULTIPOLYGON (((-87.62061 4*
Block Group 4; Census Tract 102.01; Cook County; Illi...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.68049 4*

Showing 1 to 14 of 2,162 entries, 21 total columns

estimate is the statistical estimate of median household income per block group

moe is the measure of error in the estimate

Note that 2162 rows as expected (# block groups in CCA/Chicago)

Now I can join this table with the ACS Household Income Data

NAME	variable	estimate	moe	geometry
Block Group 2; Census Tract 4005; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.62556 4*
Block Group 6; Census Tract 4804; Cook County; Illin...	B29004_001	100365	49442	*MULTIPOLYGON (((-87.58553 4*
Block Group 7; Census Tract 4804; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.58557 4*
Block Group 4; Census Tract 608; Cook County; Illinois	B29004_001	73356	7025	*MULTIPOLYGON (((-87.64794 4*
Block Group 1; Census Tract 2909; Cook County; Illin...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.73006 4*
Block Group 2; Census Tract 2909; Cook County; Illin...	B29004_001	60072	28786	*MULTIPOLYGON (((-87.72995 4*
Block Group 3; Census Tract 2909; Cook County; Illin...	B29004_001	60833	35237	*MULTIPOLYGON (((-87.73729 4*
Block Group 4; Census Tract 2909; Cook County; Illin...	B29004_001	43490	22556	*MULTIPOLYGON (((-87.73993 4*
Block Group 2; Census Tract 8387; Cook County; Illin...	B29004_001	45139	29580	*MULTIPOLYGON (((-87.72506 4*
Block Group 1; Census Tract 3102; Cook County; Illin...	B29004_001	120263	22156	*MULTIPOLYGON (((-87.64539 4*
Block Group 2; Census Tract 3102; Cook County; Illin...	B29004_001	131250	39622	*MULTIPOLYGON (((-87.64671 4*
Block Group 2; Census Tract 3201.01; Cook County; I...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.62086 4*
Block Group 1; Census Tract 3201.01; Cook County; I...	B29004_001	158651	64476	*MULTIPOLYGON (((-87.62061 4*
Block Group 4; Census Tract 102.01; Cook County; Illi...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.68049 4*

Showing 1 to 14 of 2,162 entries, 21 total columns

Issue – Some data is missing. Why?

Filter out missing data (NAs)

```
hhi_ccaf <- filter (hhi_cca, !is.na(estimate))
```

hhi_cca22 | hhi_ccaf | acs_2022variables | acs_2023variables | hhi_cca | acs_202

Filter

_1	NAME	variable	estimate	moe	geometry
ghts	Block Group 6; Census Tract 4804; Cook County; Illin...	B29004_001	100365	49442	MULTIPOLYGON (((-87.
	Block Group 4; Census Tract 608; Cook County; Illinois	B29004_001	73356	7025	MULTIPOLYGON (((-87.
dale	Block Group 2; Census Tract 2909; Cook County; Illin...	B29004_001	60072	28786	MULTIPOLYGON (((-87.
dale	Block Group 3; Census Tract 2909; Cook County; Illin...	B29004_001	60833	35237	MULTIPOLYGON (((-87.
dale	Block Group 4; Census Tract 2909; Cook County; Illin...	B29004_001	43490	22556	MULTIPOLYGON (((-87.
dale	Block Group 2; Census Tract 8387; Cook County; Illin...	B29004_001	45139	29580	MULTIPOLYGON (((-87.
Side	Block Group 1; Census Tract 3102; Cook County; Illin...	B29004_001	120263	22156	MULTIPOLYGON (((-87.
Side	Block Group 2; Census Tract 3102; Cook County; Illin...	B29004_001	131250	39622	MULTIPOLYGON (((-87.
	Block Group 1; Census Tract 3201.01; Cook County; I...	B29004_001	158651	64476	MULTIPOLYGON (((-87.
	Block Group 6; Census Tract 609; Cook County; Illinois	B29004_001	96493	30272	MULTIPOLYGON (((-87.
	Block Group 3; Census Tract 102.01; Cook County; Illi...	B29004_001	47569	15883	MULTIPOLYGON (((-87.
	Block Group 4; Census Tract 609; Cook County; Illinois	B29004_001	30450	29362	MULTIPOLYGON (((-87.
Side	Block Group 1; Census Tract 3301.03; Cook County; I...	B29004_001	94323	66240	MULTIPOLYGON (((-87.
Side	Block Group 3; Census Tract 3301.02; Cook County; I...	B29004_001	115931	41792	MULTIPOLYGON (((-87.

Showing 1 to 14 of 1,812 entries, 21 total columns

350 rows filtered out
16.19%

Seems like a lot – let's look at the previous year

ACS 2022

	variable	estimate	moe	geometry
Group 2; Census Tract 4005; Cook County; Illin...	B29004_001	54250	66162	*MULTIPOLYGON (((-87.625*
Group 6; Census Tract 4804; Cook County; Illin...	B29004_001	102770	27891	*MULTIPOLYGON (((-87.585*
Group 7; Census Tract 4804; Cook County; Illin...	B29004_001	61685	55380	*MULTIPOLYGON (((-87.585*
Group 4; Census Tract 608; Cook County; Illinois	B29004_001	71411	8511	*MULTIPOLYGON (((-87.647*
Group 1; Census Tract 2	B29004_001	16625	25018	*MULTIPOLYGON (((-87.730*
Group 2; Census Tract 2909; Cook County; Illin...	B29004_001	21708	17874	*MULTIPOLYGON (((-87.729*
Group 3; Census Tract 2909; Cook County; Illin...	B29004_001	46992	33706	*MULTIPOLYGON (((-87.737*
Group 4; Census Tract 2909; Cook County; Illin...	B29004_001	48237	13084	*MULTIPOLYGON (((-87.739*
Group 2; Census Tract 8387; Cook County; Illin...	B29004_001	36250	16529	*MULTIPOLYGON (((-87.725*
Group 1; Census Tract 3102; Cook County; Illin...	B29004_001	121012	26103	*MULTIPOLYGON (((-87.645*
Group 2; Census Tract 3102; Cook County; Illin...	B29004_001	136182	43751	*MULTIPOLYGON (((-87.646*
Group 2; Census Tract 3201.01; Cook County; I...	B29004_001	*NA*	*NA*	*MULTIPOLYGON (((-87.620*
Group 1; Census Tract 3201.01; Cook County; I...	B29004_001	127273	93131	*MULTIPOLYGON (((-87.620*
Group 4; Census Tract 102.01; Cook County; Illi...	B29004_001	37367	38907	*MULTIPOLYGON (((-87.680*

Block Group 4; Census Tract 608; Cook County; Illinois

Showing 1 to 14 of 2,162 entries, 21 total columns

Only 11 rows filtered out
.51%

Filtered out NAs

	variable	estimate	moe	geometry
Group 2; Census Tract 4005; Cook County; Illin...	B29004_001	54250	66162	*MULTIPOLYGON (((-87.62*
Group 6; Census Tract 4804; Cook County; Illin...	B29004_001	102770	27891	*MULTIPOLYGON (((-87.58*
Group 7; Census Tract 4804; Cook County; Illin...	B29004_001	61685	55380	*MULTIPOLYGON (((-87.58*
Group 4; Census Tract 608; Cook County; Illinois	B29004_001	71411	8511	*MULTIPOLYGON (((-87.64*
Group 1; Census Tract 2909; Cook County; Illin...	B29004_001	16625	2 8511	*MULTIPOLYGON (((-87.73*
Group 2; Census Tract 2909; Cook County; Illin...	B29004_001	21708	17874	*MULTIPOLYGON (((-87.72*
Group 3; Census Tract 2909; Cook County; Illin...	B29004_001	46992	33706	*MULTIPOLYGON (((-87.73*
Group 4; Census Tract 2909; Cook County; Illin...	B29004_001	48237	13084	*MULTIPOLYGON (((-87.73*
Group 2; Census Tract 8387; Cook County; Illin...	B29004_001	36250	16529	*MULTIPOLYGON (((-87.72*
Group 1; Census Tract 3102; Cook County; Illin...	B29004_001	121012	26103	*MULTIPOLYGON (((-87.64*
Group 2; Census Tract 3102; Cook County; Illin...	B29004_001	136182	43751	*MULTIPOLYGON (((-87.64*
Group 1; Census Tract 3201.01; Cook County; I...	B29004_001	127273	93131	*MULTIPOLYGON (((-87.62*
Group 4; Census Tract 102.01; Cook County; Illi...	B29004_001	37367	38907	*MULTIPOLYGON (((-87.68*
Group 6; Census Tract 609; Cook County; Illinois	B29004_001	96288	32009	*MULTIPOLYGON (((-87.64*

Showing 1 to 14 of 2,151 entries, 21 total columns

Much better!
Use 2022 instead of 2023

Is 2021 even better?

	variable	estimate	moe	geometry
roup 2, Census Tract 4005, Cook County, Illin...	B29004_001	21148	12727	*MULTIPOLYGON (*
roup 6, Census Tract 4804, Cook County, Illin...	B29004_001	95944	36971	*MULTIPOLYGON (*
roup 7, Census Tract 4804, Cook County, Illin...	B29004_001	33805	7519	*MULTIPOLYGON (*
roup 4, Census Tract 608, Cook County, Illinois	B29004_001	70682	9420	*MULTIPOLYGON (*
roup 1, Census Tract 2909, Cook County, Illin...	B29004_001	16694	16206	*MULTIPOLYGON (*
roup 2, Census Tract 2909, Cook County, Illin...	B29004_001	26602	50694	*MULTIPOLYGON (*
roup 3, Census Tract 2909, Cook County, Illin...	B29004_001	31681	14119	*MULTIPOLYGON (*
roup 4, Census Tract 2909, Cook County, Illin...	B29004_001	43271	17726	*MULTIPOLYGON (*
roup 2, Census Tract 8387, Cook County, Illin...	B29004_001	40417	18869	*MULTIPOLYGON (*
roup 1, Census Tract 3102, Cook County, Illin...	B2900	96875	39089	*MULTIPOLYGON (*
roup 2, Census Tract 3102, Cook County, Illin...	B29004_001	92174	45357	*MULTIPOLYGON (*
roup 1, Census Tract 3201.01, Cook County, I...	B29004_001	138939	123568	*MULTIPOLYGON (*
roup 4, Census Tract 102.01, Cook County, Illi...	B29004_001	37981	59758	*MULTIPOLYGON (*
roup 6, Census Tract 609, Cook County, Illinois	B29004_001	110345	26541	*MULTIPOLYGON (*

B29004_001

Showing 1 to 14 of 2,153 entries, 21 total columns

Only 9 rows filtered out
.42%

Tradeoff
Small gain in data volume
vs timeliness of data

What to do?

Community Analysis (Ch 6 in Social Media Mining)
Social Community

[real-world] community

A group of individuals with common *economic*, *social*, or *political* interests or characteristics, often living in *relative proximity*.

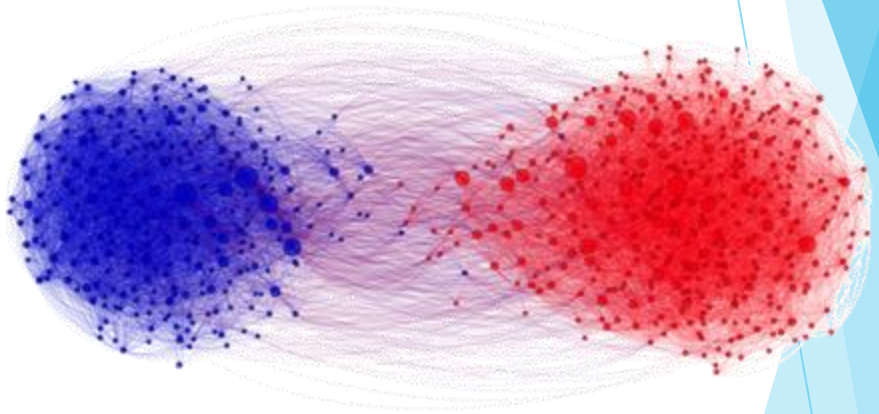
Why analyze communities?

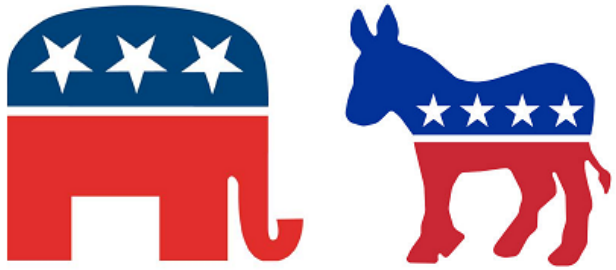
Analyzing communities helps better understand users

- Users form groups based on their interests

Groups provide a clear global view of user interactions

- E.g., find polarization





Some behaviors are only observable in a group setting and not on an individual level

- Some republican can **agree** with some democrats, but their parties can **disagree**

Social Media Communities

- **Formation:**
  - When like-minded users on social media form a link and start interacting with each other
- **More Formal Formation**:
  1. A set of at least two nodes sharing some interest, and
  2. Interactions with respect to that interest.
- Social Media Communities
  - **Explicit** (*emic*): formed by user subscriptions
  - **Implicit** (*etic*): implicitly formed by social interactions
    - **Example:** individuals calling Canada from the United States
    - Phone operator considers them one community for promotional offers
- Other community names: *group*, *cluster*, *cohesive subgroup*, or *module*

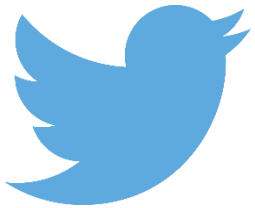
Explicit (emic) Community

- Satisfy these criteria
  - Community members understand they are its members
  - Nonmembers understand who the community members are
  - Community members often have more interactions with each other than with nonmembers

Examples of Explicit Social Media Communities

Facebook has groups and communities. Users can

- post messages and images,
- can comment on other messages,
- can like posts, and
- can view activities of other users



In X (Twitter), communities form as lists.

– Users join lists to receive information in the form of tweets

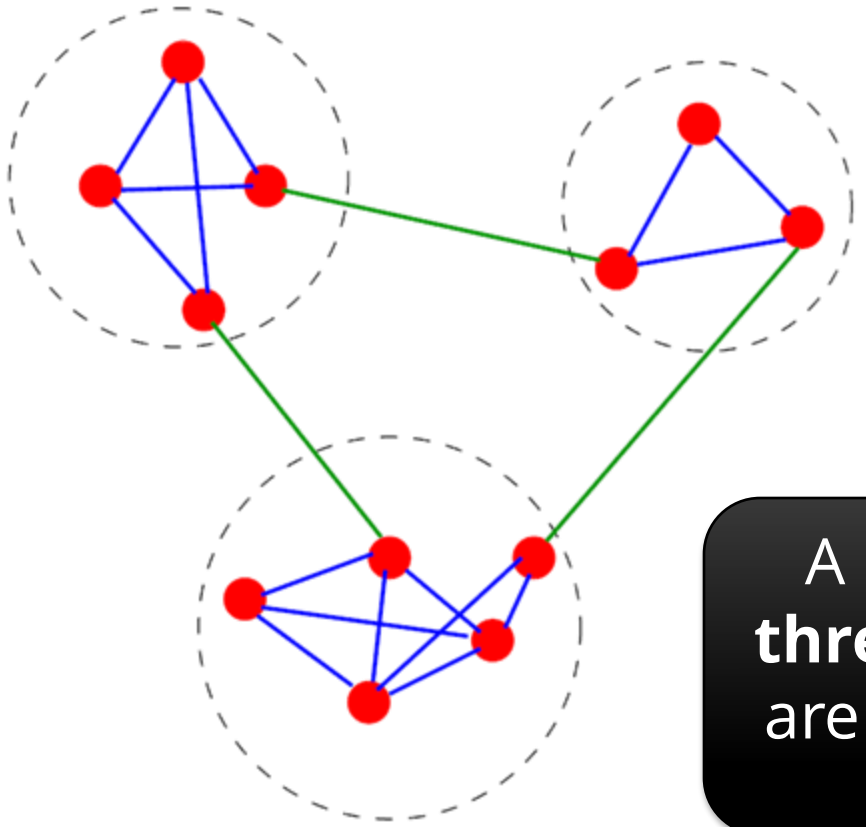
LinkedIn provides *Groups* and *Associations*.

– Users can join professional groups where they can post and share information related to the group

Implicit (etic) Communities

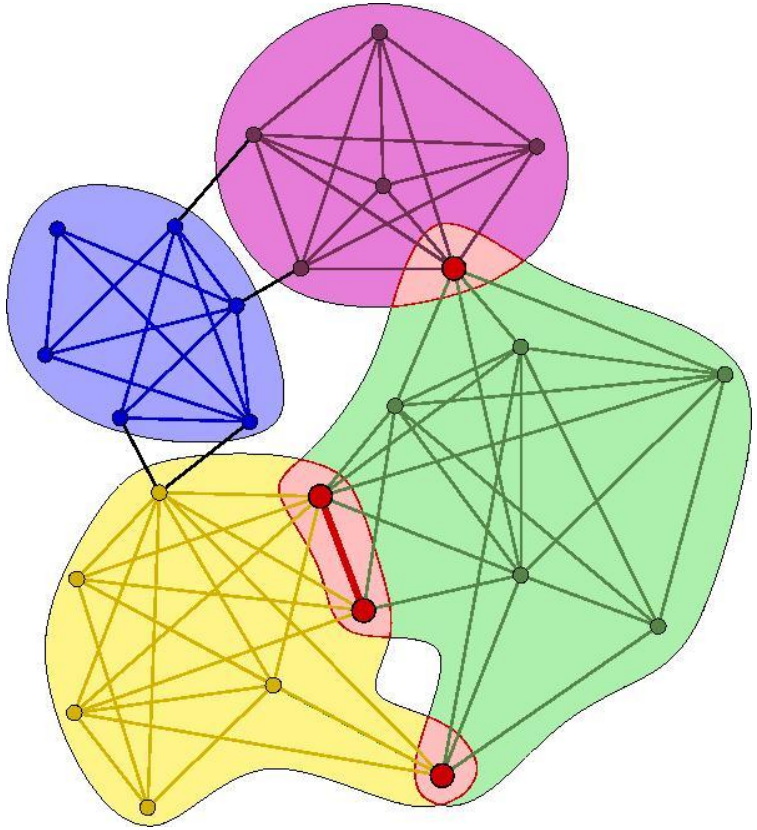
- Individuals tacitly interact with others in the form of an unacknowledged community
- This chapter focuses on finding implicit communities in social media

Finding Implicit Communities: An Example

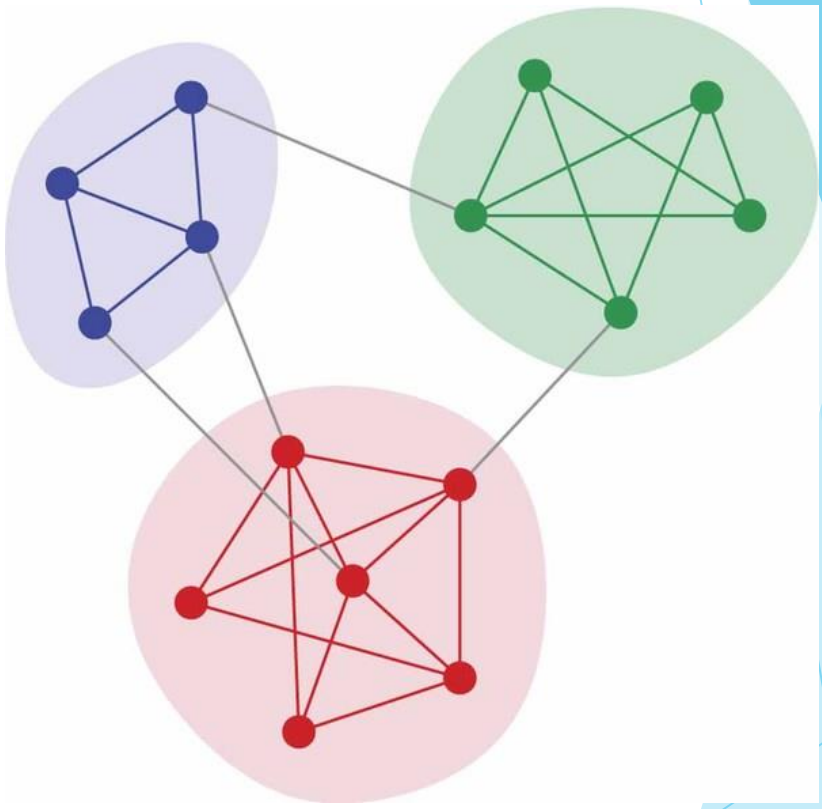


A simple graph in which **three** implicit communities are found, enclosed by the dashed circles

Overlapping vs. Disjoint Communities



Overlapping Communities

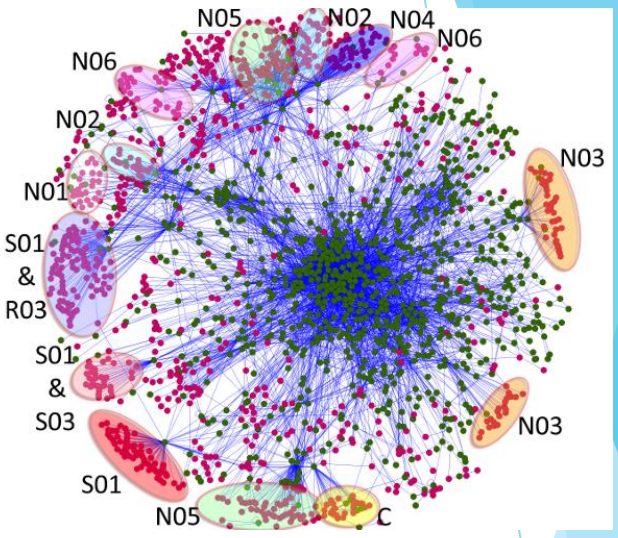


Disjoint Communities

Implicit communities in other domains

Protein-protein interaction networks

- Communities are likely to group proteins having the same specific function within the cell



World Wide Web

- Communities may correspond to groups of pages dealing with the same or related topics

Metabolic networks

- Communities may be related to functional modules such as cycles and pathways

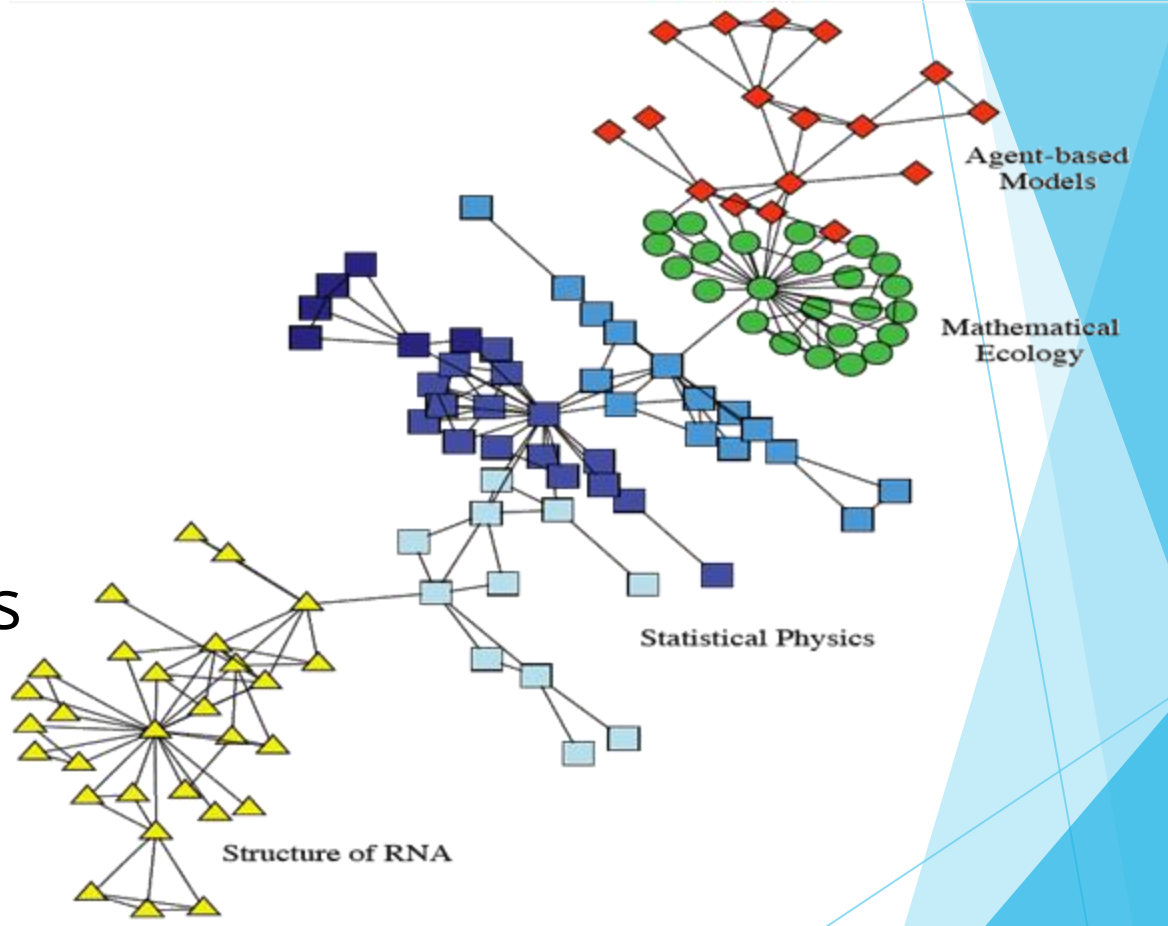
Food webs

- Communities may identify compartments

Real-world Implicit Communities

Collaboration network between scientists working at the Santa Fe Institute.

The colors indicate high level communities and correspond to research divisions of the institute



What is Community Analysis?

- **Community detection**
  - Discovering implicit communities
- **Community evolution**
  - Studying temporal evolution of communities
- **Community evaluation**
  - Evaluating Detected Communities

Community Detection

Social Science Perspective

Bottom-up or Member-Based Community Detection

- Many of the approaches to understanding the structure of a network emphasize how dense connections are built-up from simpler dyads and triads to more extended dense clusters such as "cliques."
- This view of social structure focuses attention on how solidarity and connection of large social structures can be built up out of small and tight components: a sort of "bottom up" approach.
- Network analysts have developed a number of useful definitions an algorithms that identify how larger structures are compounded from smaller ones: cliques, n-cliques, n-clans, and k-plexes all look at networks this way.

Why identify communities in social network?

- Divisions of actors into groups and sub-structures can be a very important aspect of social structure.
- It can be important in understanding how the network as a whole is likely to behave.
- Suppose the actors in one network form two non-overlapping groups; and, suppose that the actors in another network also form two groups, but that the memberships overlap (some people are members of both groups).
- Where the groups overlap, we might expect that conflict between them is less likely than when the groups don't overlap.
- Where the groups overlap, mobilization and diffusion may spread rapidly across the entire network; where the groups don't overlap, traits may occur in one group and not diffuse to the other.

https://faculty.ucr.edu/~hanneman/nettext/C11_Cliques.html

What can communities tell us about the individuals that are part of it?

- Knowing how an individual is embedded in the structure of groups within a net may also be critical to understanding his/her behavior.
  - For example, some people may act as "bridges" between groups (cosmopolitans, boundary spanners, or "brokers").
  - Others may have all of their relationships within a single group (locals or insiders).
  - Some actors may be part of a tightly connected and closed elite, while others are completely isolated from this group.
  - Such differences in the ways that individuals are embedded in the structure of groups within in a network can have profound consequences for the ways that these actors see their "society," and the behaviors that they are likely to practice.

https://faculty.ucr.edu/~hanneman/nettext/C11_Cliques.html

Top-down or group-based community detection

- Looking at the whole network, we can think of sub-structures as areas of the graph that seem to be locally dense, but separated to some degree, from the rest of the graph.
  - components, blocks/cutpoints, K-cores, Lambda sets and bridges, factions, and f-groups
- The idea that some regions of a graph may less connected to the whole than others may lead to insights into lines of cleavage and division.
- Weaker parts in the "social fabric" also create opportunities for brokerage and less constrained action.
- So, the numbers and sizes of regions, and their "connection topology" may be consequential for predicting both the opportunities and constraints facing groups and actors, as well as predicting the evolution of the graph itself.

https://faculty.ucr.edu/~hanneman/nettext/C11_Cliques.html

Some key questions about a graph

- How separate are the sub-graphs? Do they overlap and share members, or do they divide or factionalize the network?
- How large are the connected sub-graphs? Are there a few big groups, or a larger number of small groups?
- Are there particular actors that appear to play network roles? For example, act as nodes that connect the graph, or who are isolated from groups?

https://faculty.ucr.edu/~hanneman/nettext/C11_Cliques.html

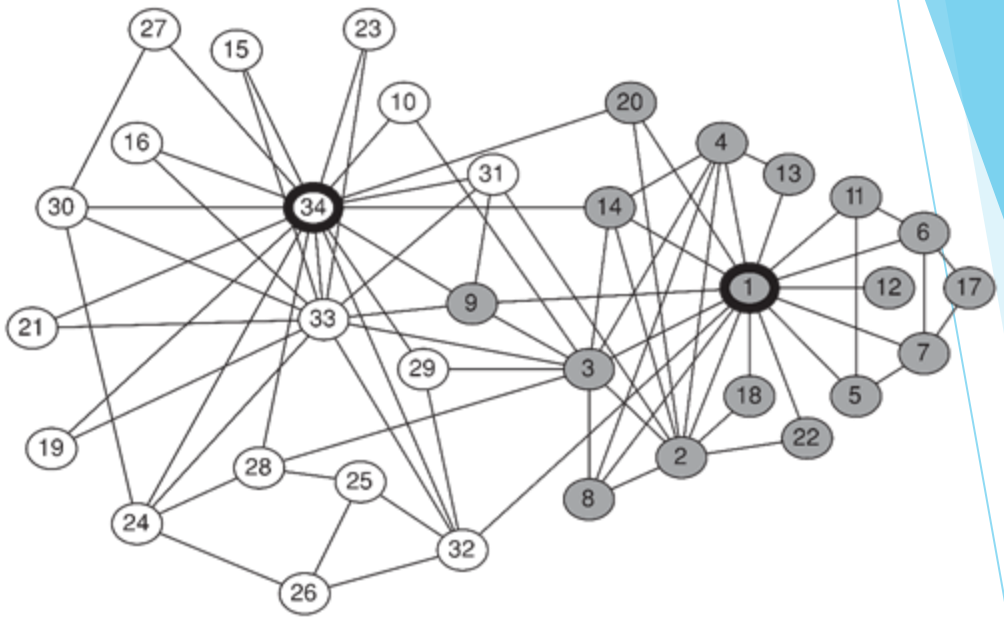
What is community detection?

- The process of finding clusters of nodes ("*communities*")
  - With **Strong** internal connections and
  - **Weak** connections between different communities
- Ideal decomposition of a large graph
  - Completely disjoint communities
  - There are no interactions between different communities.
- In practice,
  - find community partitions that are maximally decoupled.

Why Detecting Communities is Important?

Zachary's karate club

Interactions between
34 members of a
karate club
for over two years



- The club members split into two groups (**gray** and **white**)
- Disagreement between the administrator of the club (node **34**) and the club's instructor (node **1**),
- The members of one group left to start their own club

The same communities can be found using community detection

Why Community Detection?

Network Summarization

- A community can be considered as a summary of the whole network
- Easier to visualize and understand

Preserve Privacy

- [Sometimes] a community can reveal some properties without releasing the individuals' privacy information.

Community Detection vs. Clustering

Clustering

- Data is often non-linked (matrix rows)
- Clustering works on the distance or similarity matrix, e.g., k-means.
- If you use k-means with adjacency matrix rows, you are only considering the ego-centric network

Community detection

- Data is linked (a graph)
- Network data tends to be “discrete”, leading to algorithms using the graph property directly
  - k-clique, quasi-clique, or edge-betweenness