

CLIQUE PERCOLATION METHOD (CPM)

Slides
mostly by

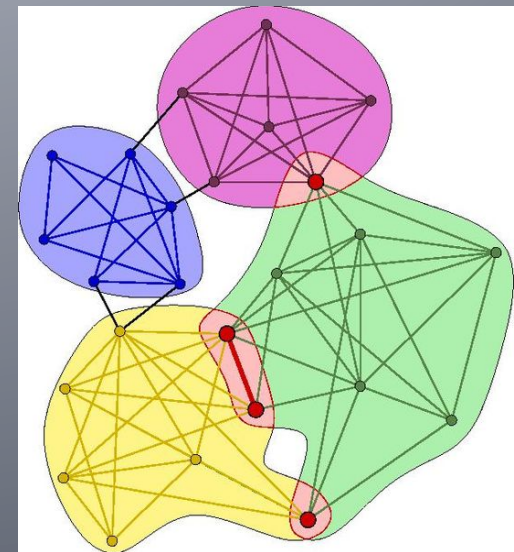
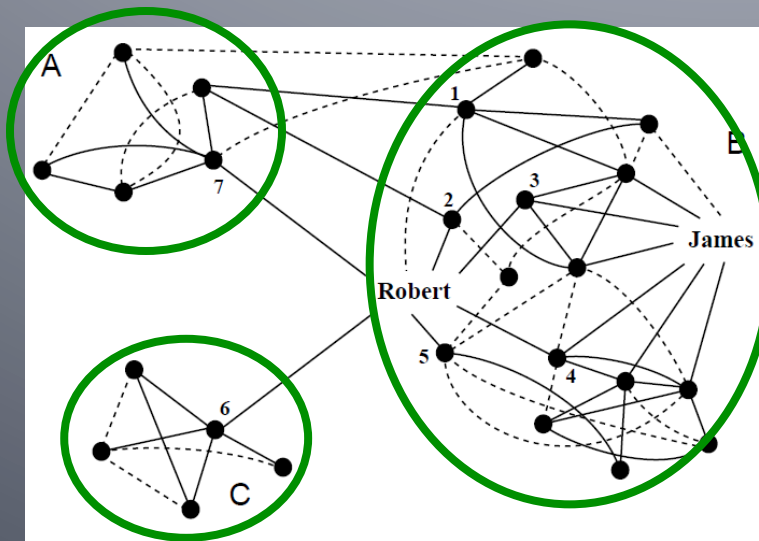
Eugene Lim

CONTENTS

- What is CPM?
- Algorithm
- Analysis
- Conclusion

Community Detection

How to find communities?



We will work with **undirected** (unweighted) networks

WHAT IS CPM?

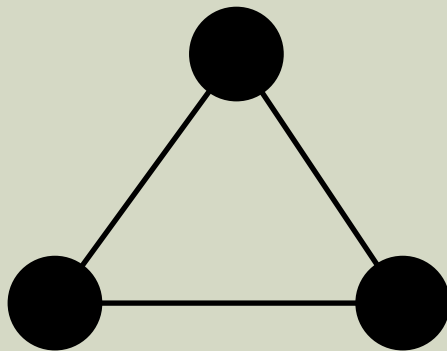
- Method to find overlapping communities
- Based on concept:
 - internal edges of community likely to form cliques
 - Intercommunity edges unlikely to form cliques

CLIQUE

- Clique: Complete graph
- k-clique: Complete graph with k vertices

CLIQUE

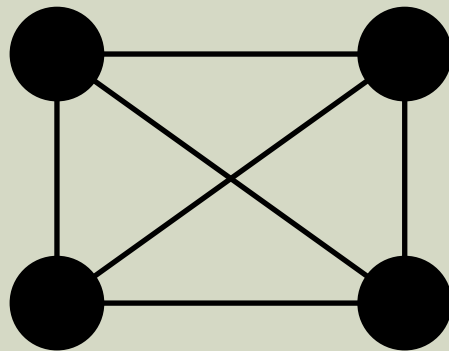
- Clique: Complete graph
- k-clique: Complete graph with k vertices



3-clique

CLIQUE

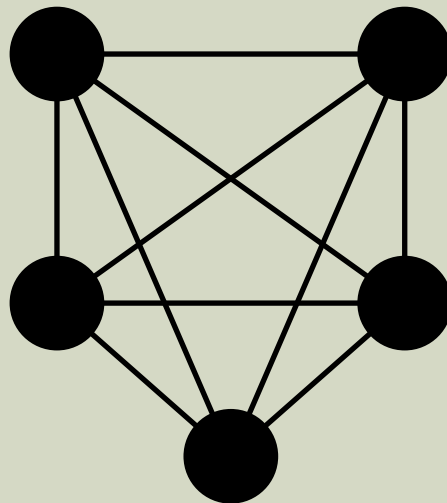
- Clique: Complete graph
- k-clique: Complete graph with k vertices



4-clique

CLIQUE

- Clique: Complete graph
- k-clique: Complete graph with k vertices



5-clique

K-CLIQUE COMMUNITIES

- Adjacent k-cliques

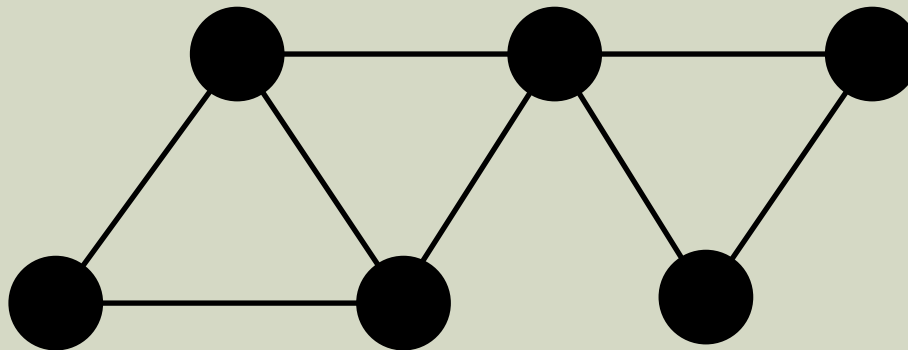
Two k-cliques are adjacent when they share $k-1$ nodes

K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$

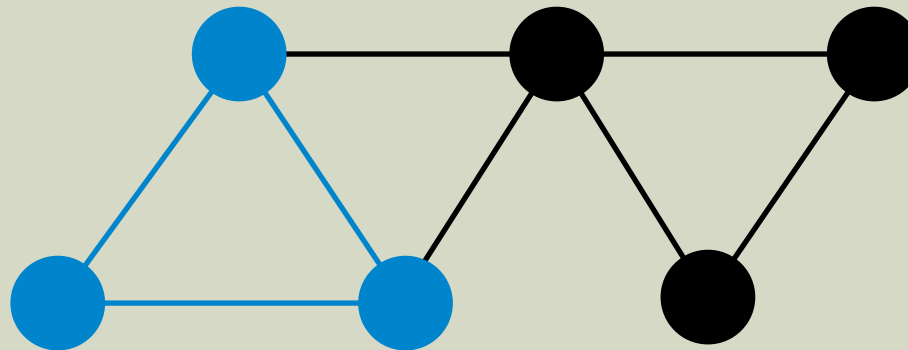


K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$



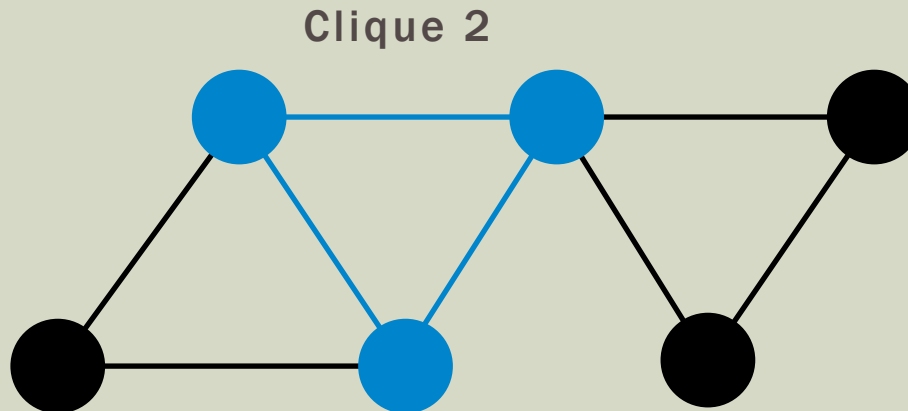
Clique 1

K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$

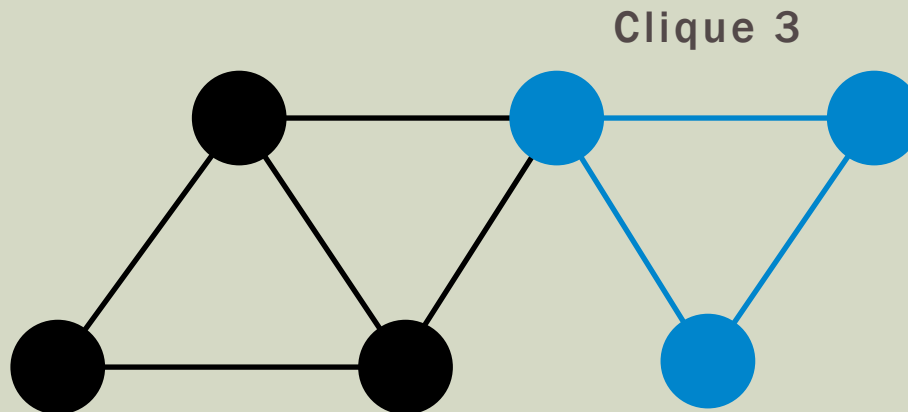


K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$

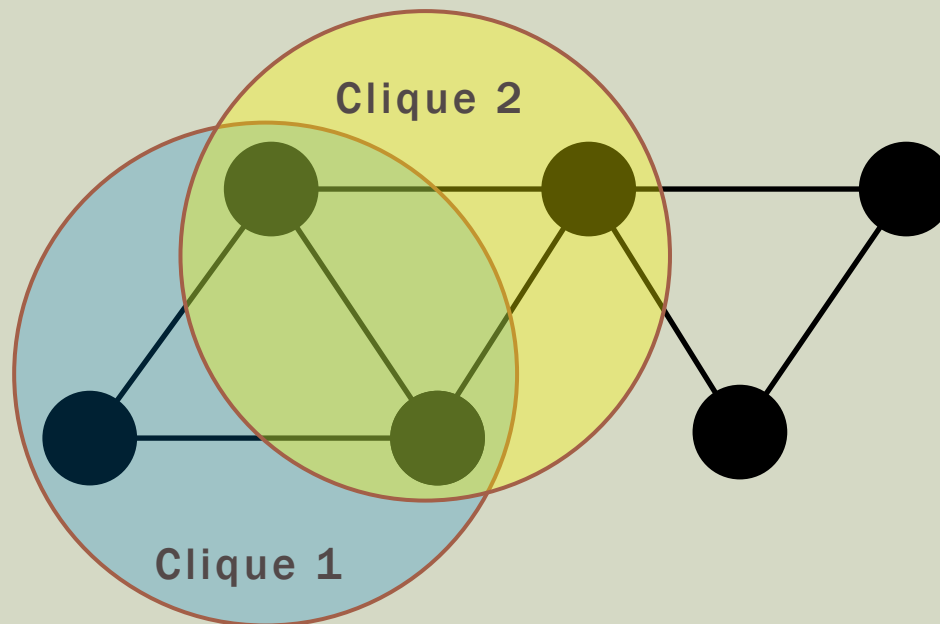


K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$

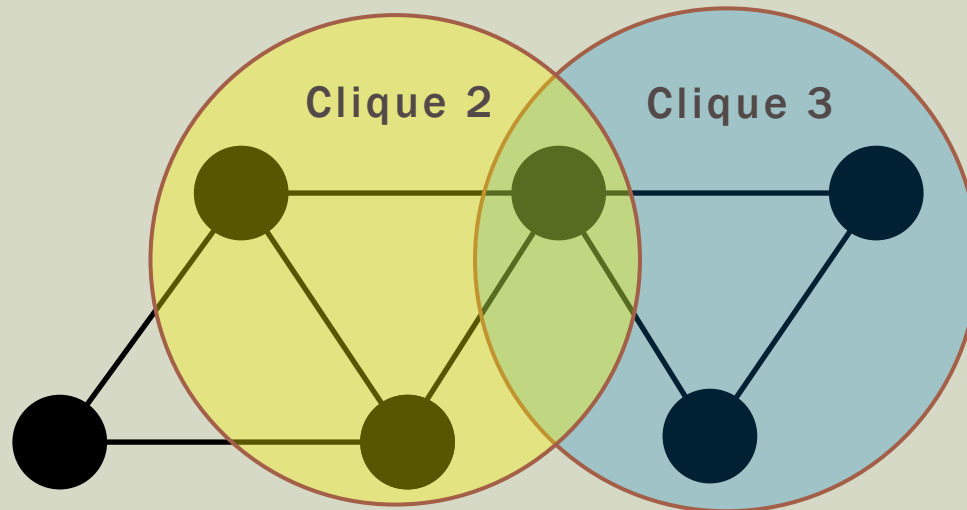


K-CLIQUE COMMUNITIES

■ Adjacent k-cliques

Two k-cliques are adjacent when they share $k-1$ nodes

$k = 3$



K-CLIQUE COMMUNITIES

- **k-clique community**

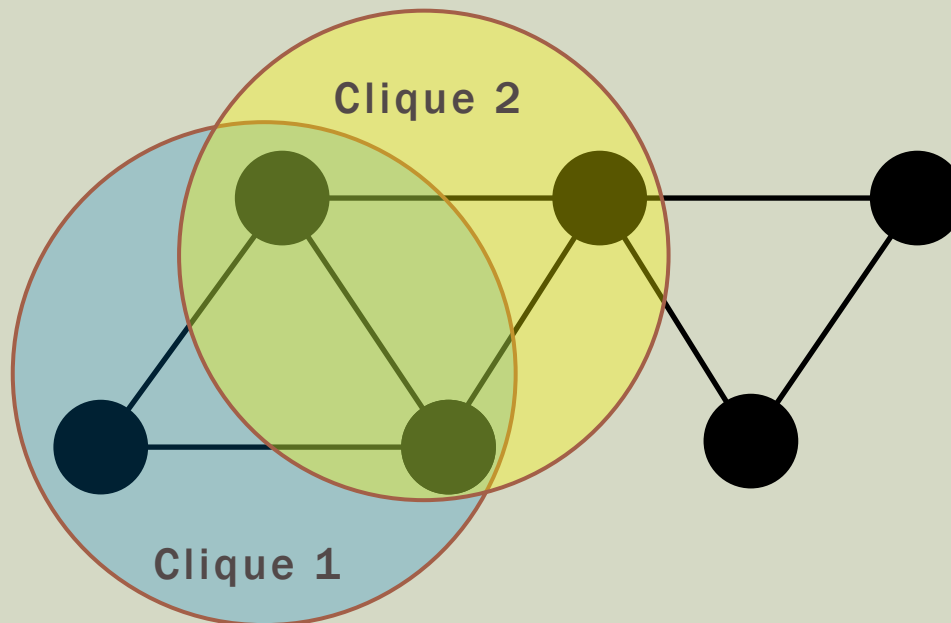
Union of all k-cliques that can be reached from each other through a series of adjacent k-cliques

K-CLIQUE COMMUNITIES

- **k-clique community**

Union of all k-cliques that can be reached from each other through a series of adjacent k-cliques

$k = 3$

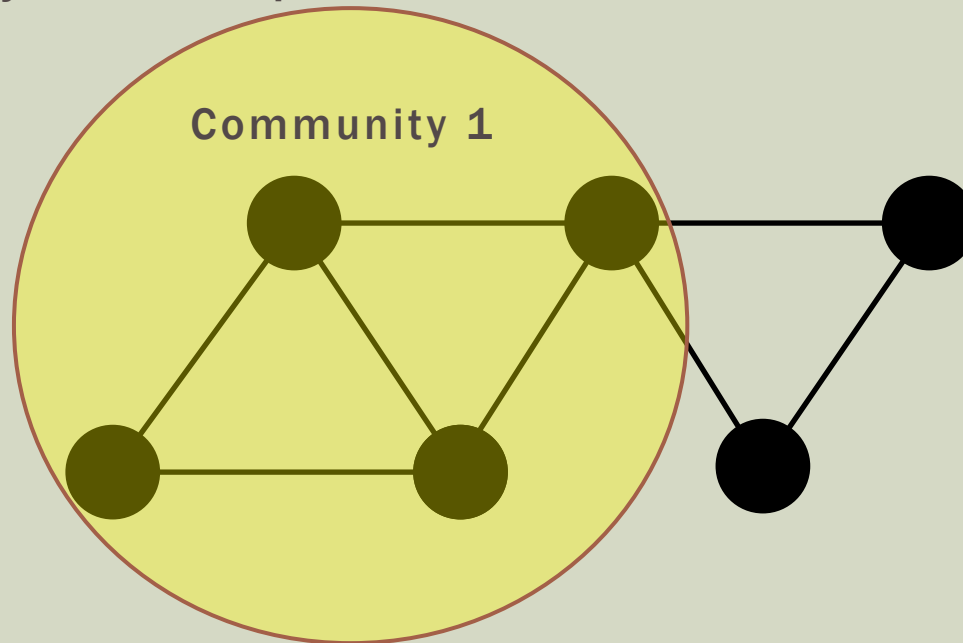


K-CLIQUE COMMUNITIES

- **k-clique community**

Union of all k-cliques that can be reached from each other through a series of adjacent k-cliques

$k = 3$

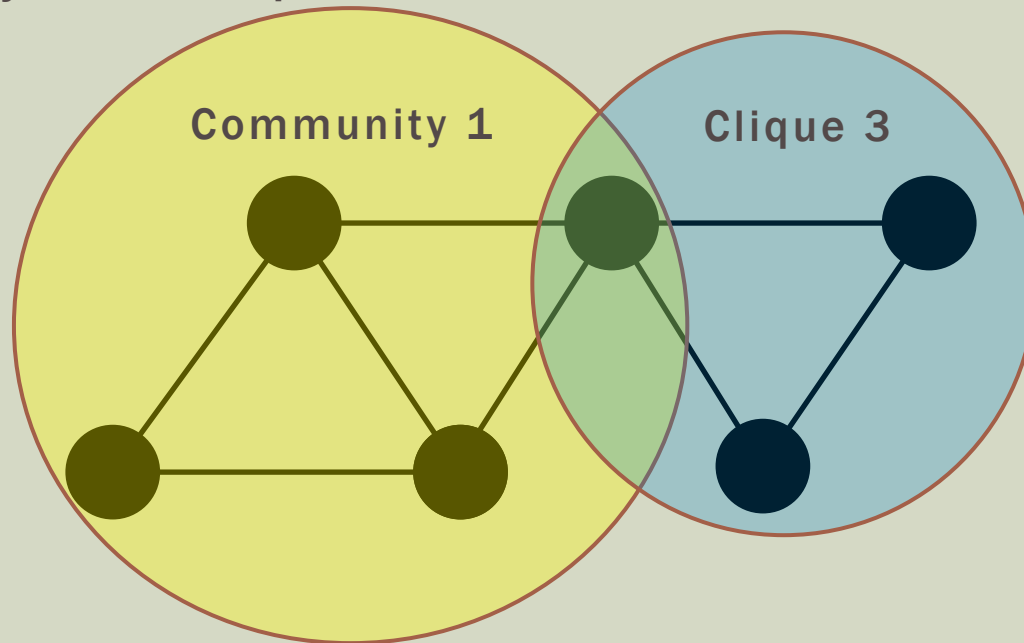


K-CLIQUE COMMUNITIES

- **k-clique community**

Union of all k-cliques that can be reached from each other through a series of adjacent k-cliques

$k = 3$

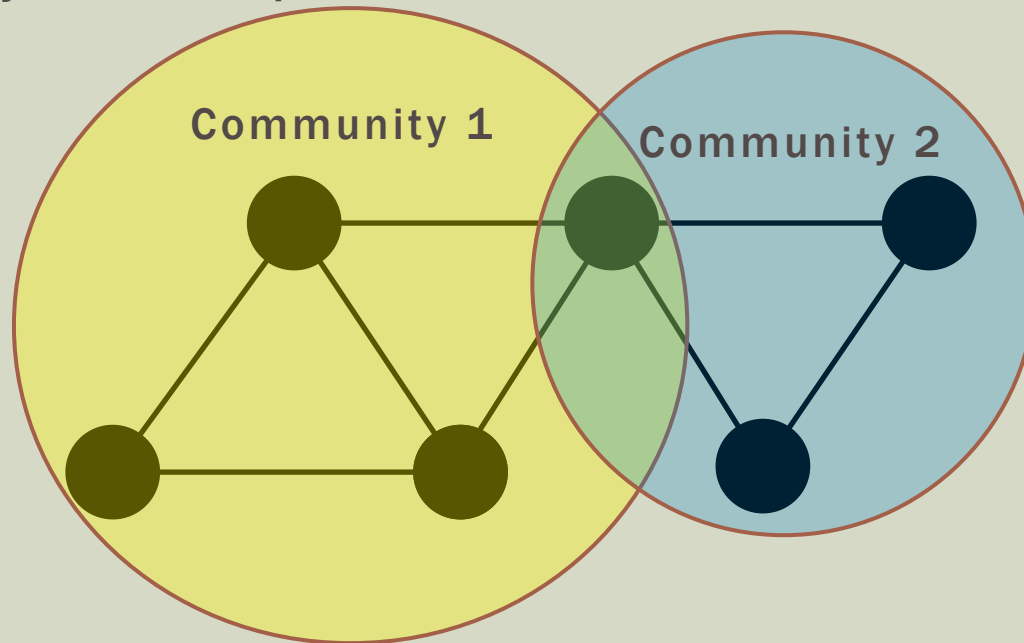


K-CLIQUE COMMUNITIES

- **k-clique community**

Union of all k-cliques that can be reached from each other through a series of adjacent k-cliques

$k = 3$



ALGORITHM

- Locate maximal cliques
- Convert from cliques to k-clique communities

LOCATE MAXIMAL CLIQUES

- Largest possible clique size can be determined from degrees of vertices
- Starting from this size, find all cliques, then reduce size by 1 and repeat

LOCATE MAXIMAL CLIQUES

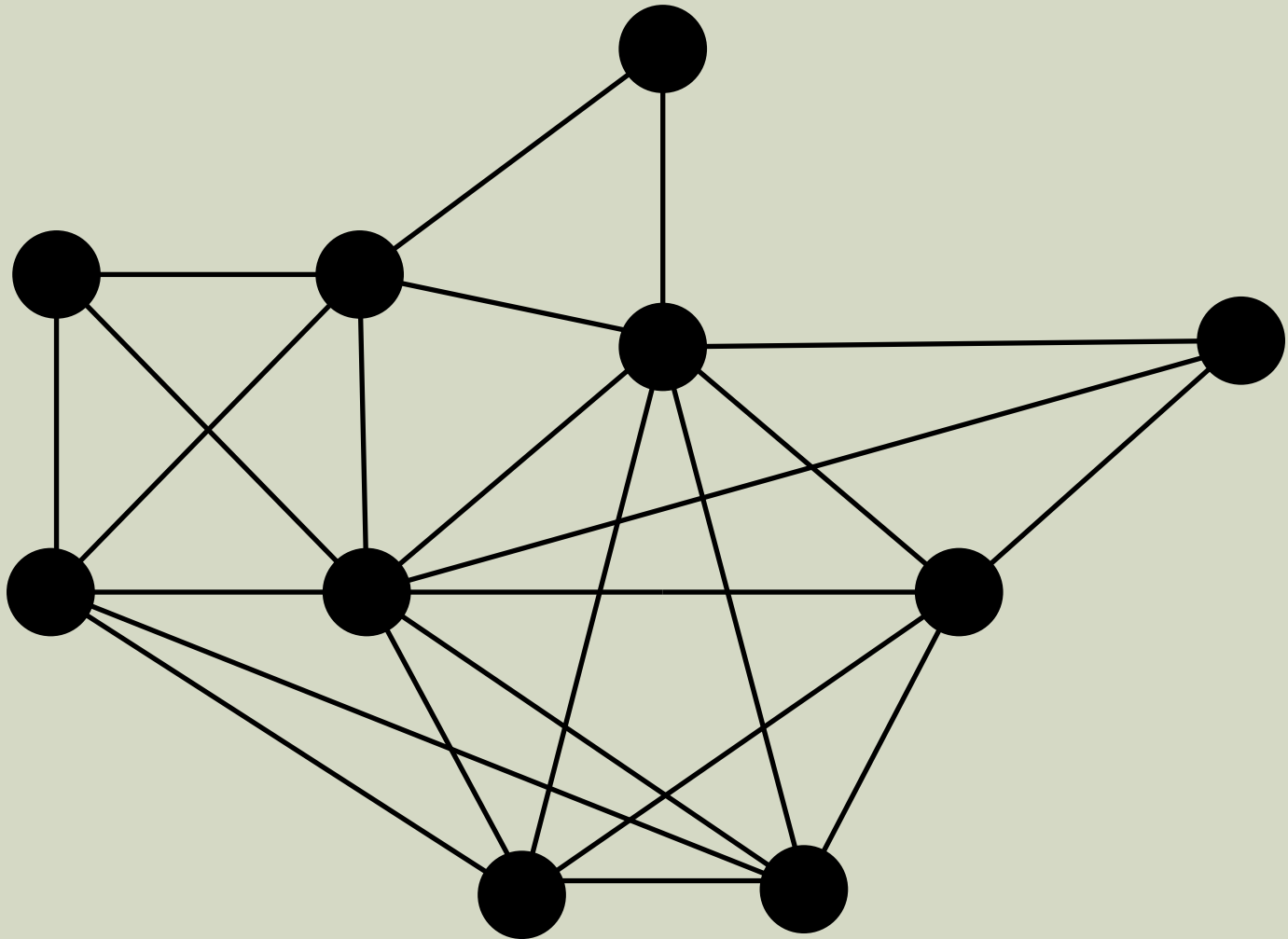
- Finding all cliques: brute-force

1. Set A initially contains vertex v , Set B contains neighbours of v
2. Transfer one vertex w from B to A
3. Remove vertices that are not neighbours of w from B
4. Repeat until A reaches desired size
5. If fail, step back and try other possibilities

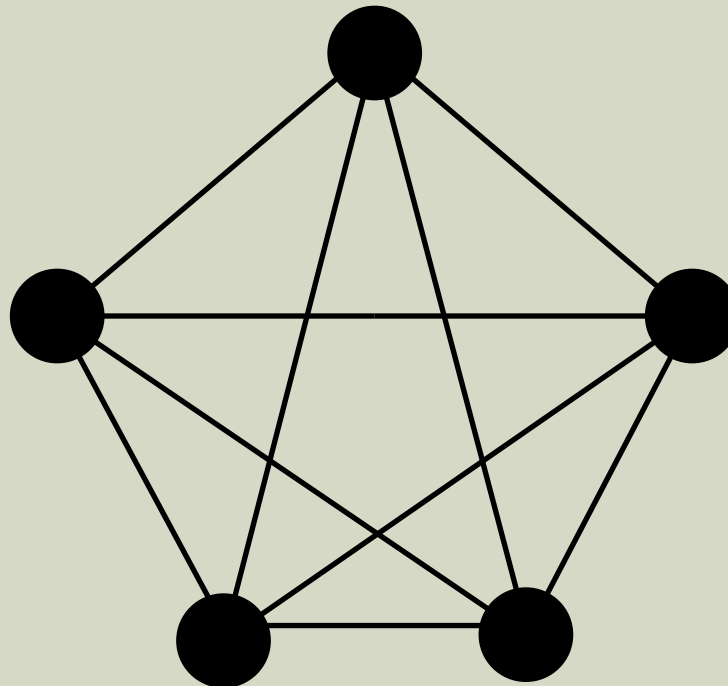
ALGORITHM

- Locate maximal cliques
- Convert from cliques to k-clique communities

CLIQUEES TO K-CLIQUE COMMUNITIES

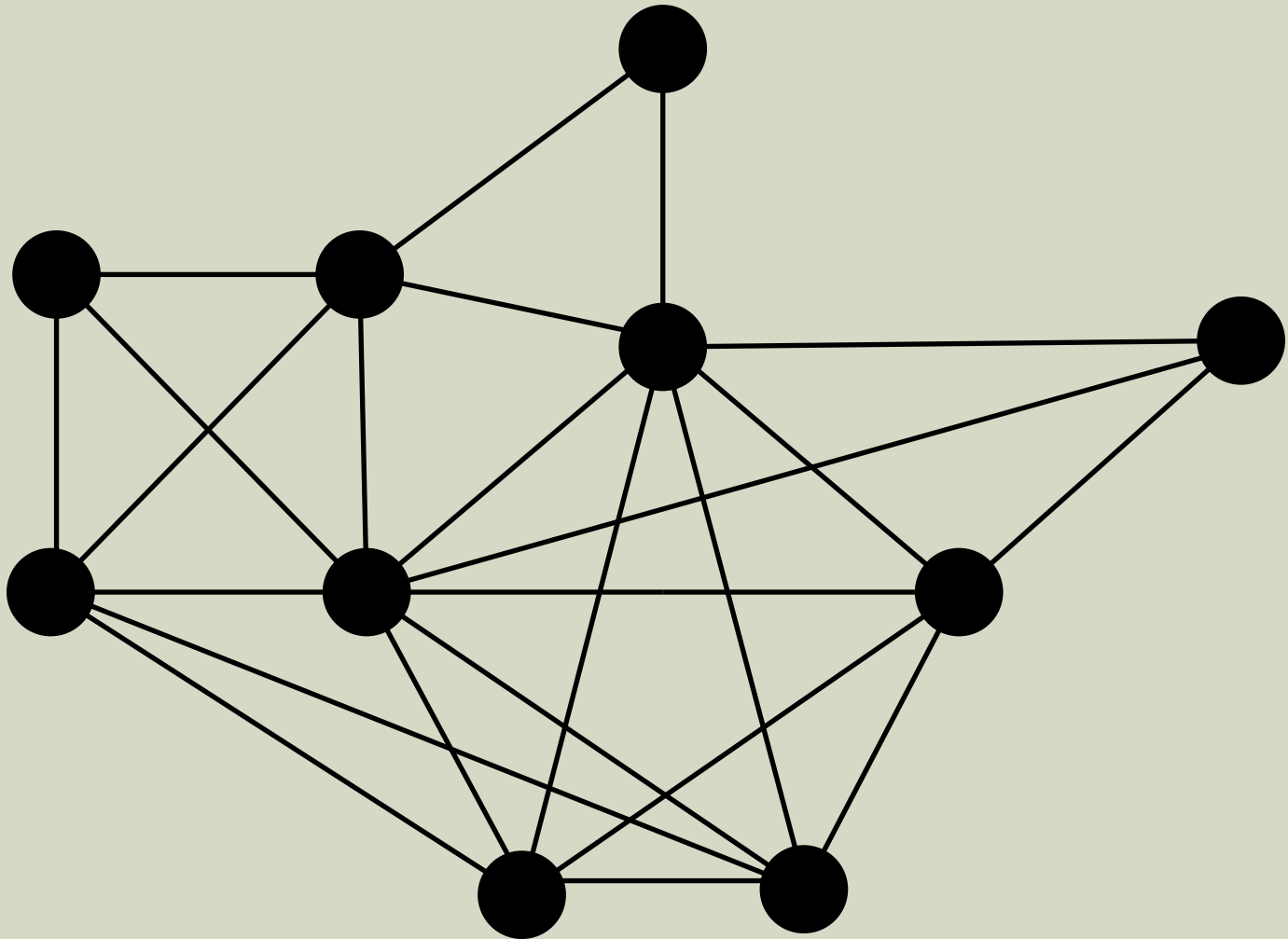


CLIQUE TO K-CLIQUE COMMUNITIES

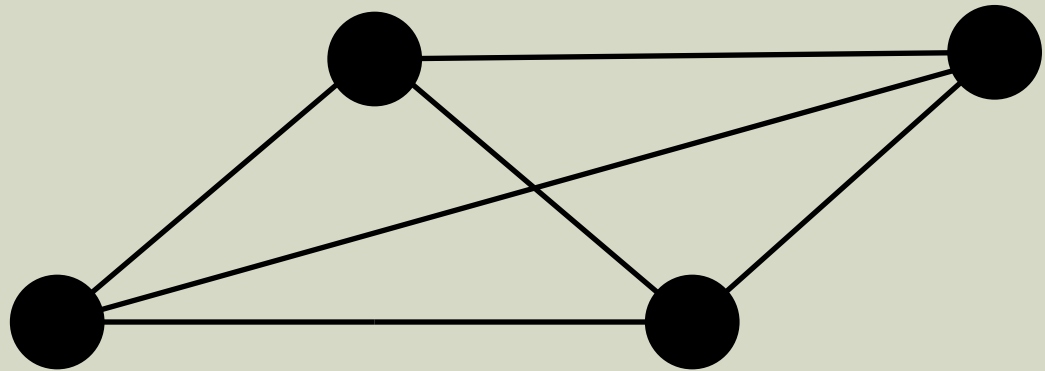


Clique 1: 5-clique

CLIQUEES TO K-CLIQUE COMMUNITIES

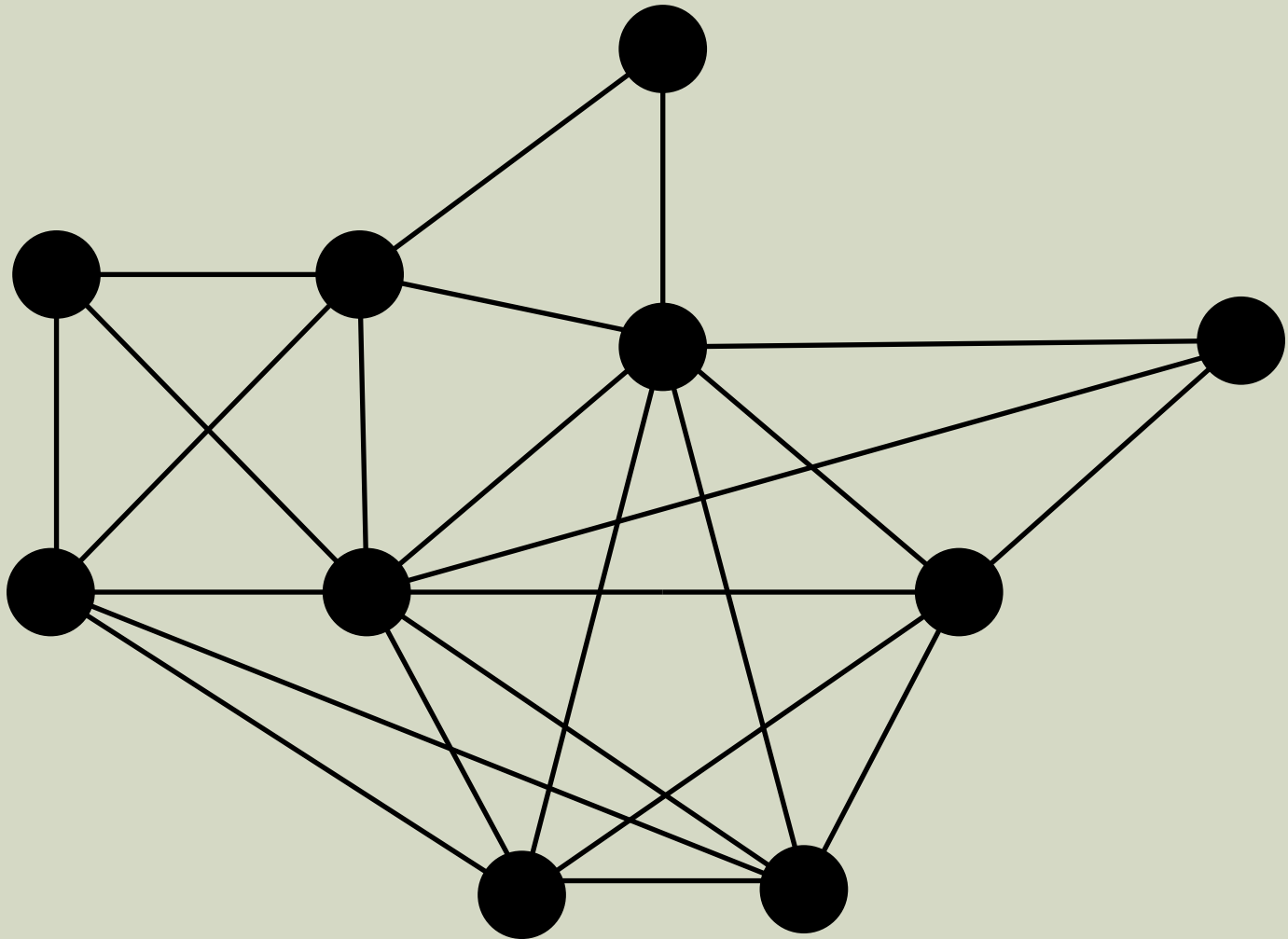


CLIQUE TO K-CLIQUE COMMUNITIES



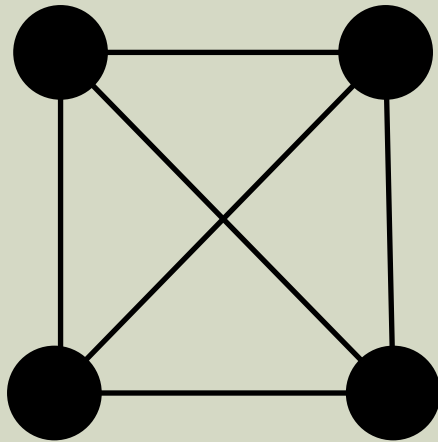
Clique 2: 4-clique

CLIQUEES TO K-CLIQUE COMMUNITIES

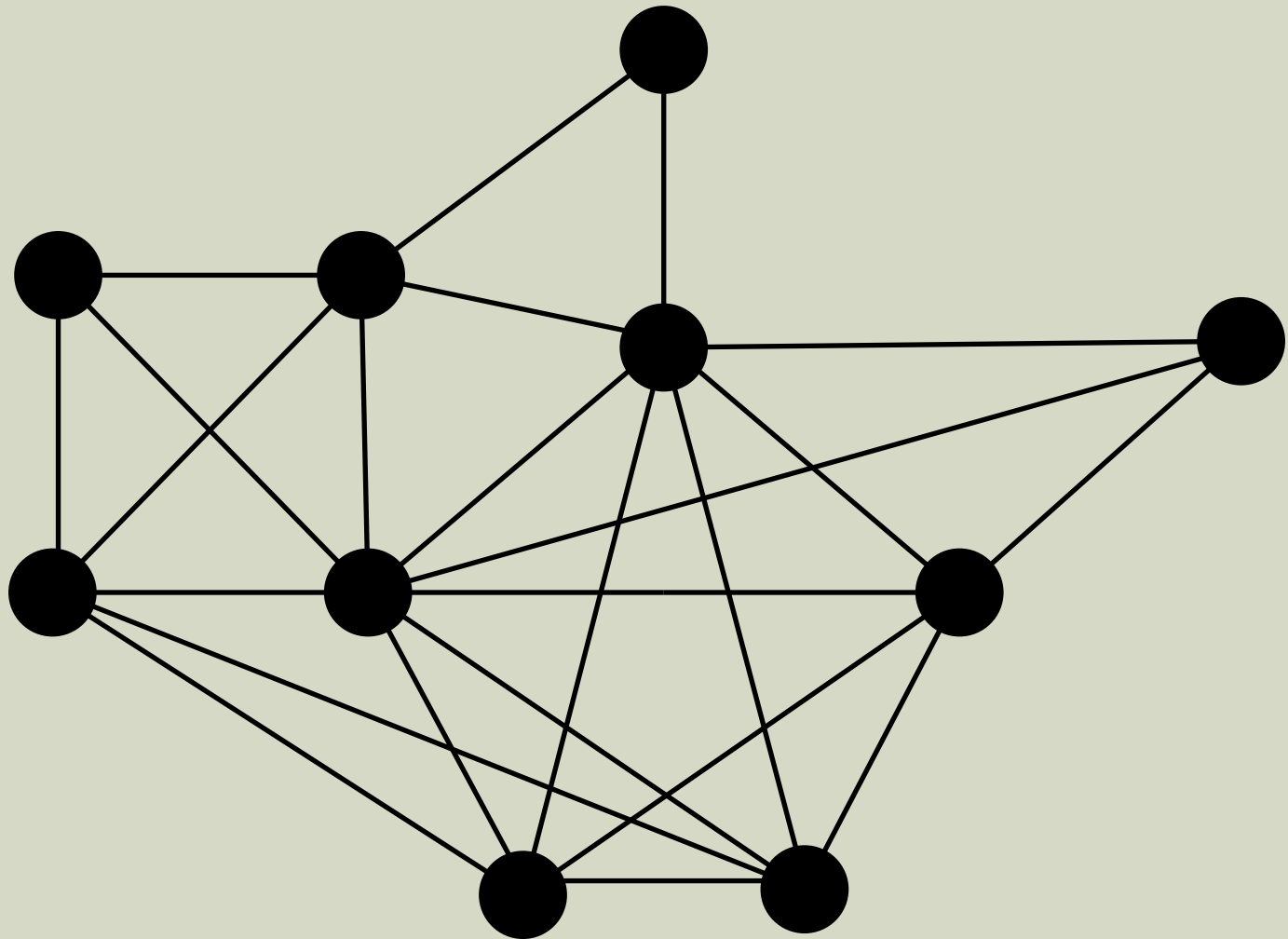


CLIQUE TO K-CLIQUE COMMUNITIES

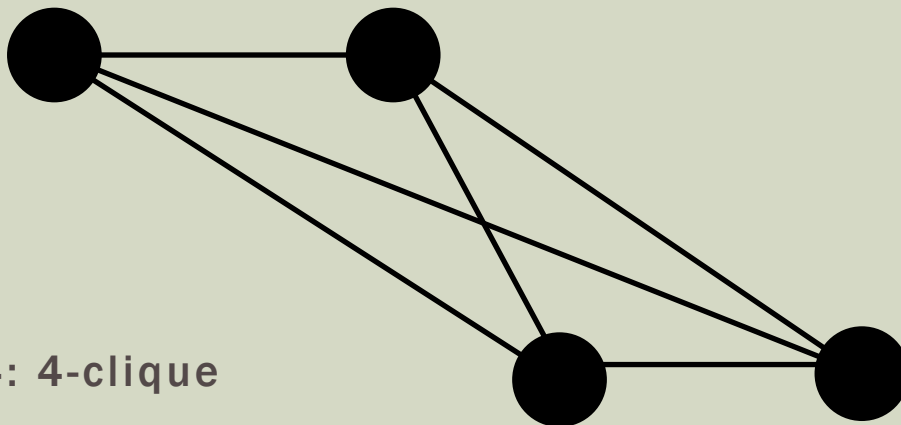
Clique 3: 4-clique



CLIQUEES TO K-CLIQUE COMMUNITIES

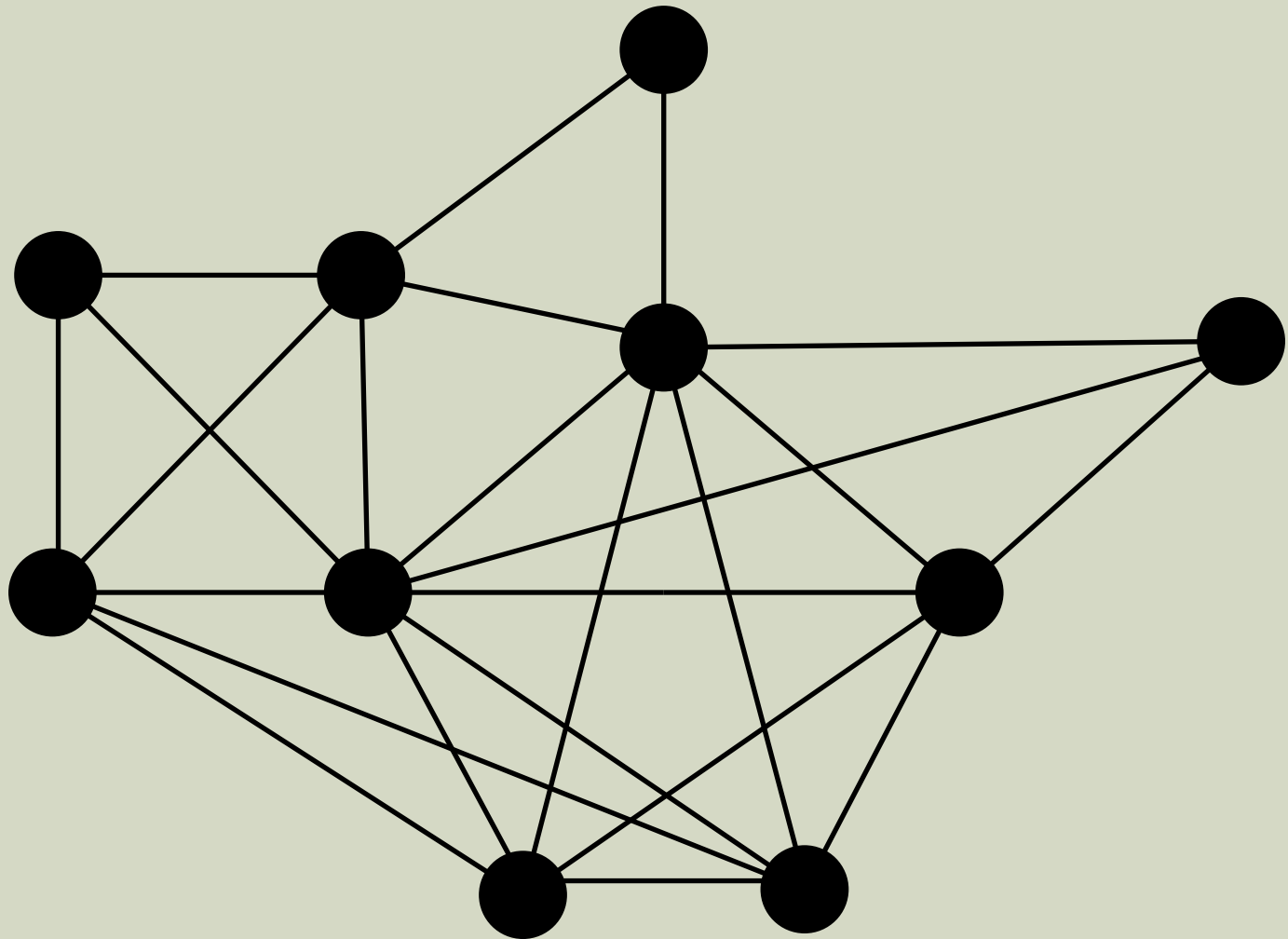


CLIQUE TO K-CLIQUE COMMUNITIES

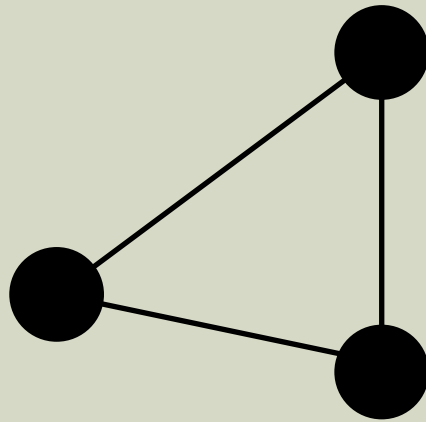


Clique 4: 4-clique

CLIQUEES TO K-CLIQUE COMMUNITIES

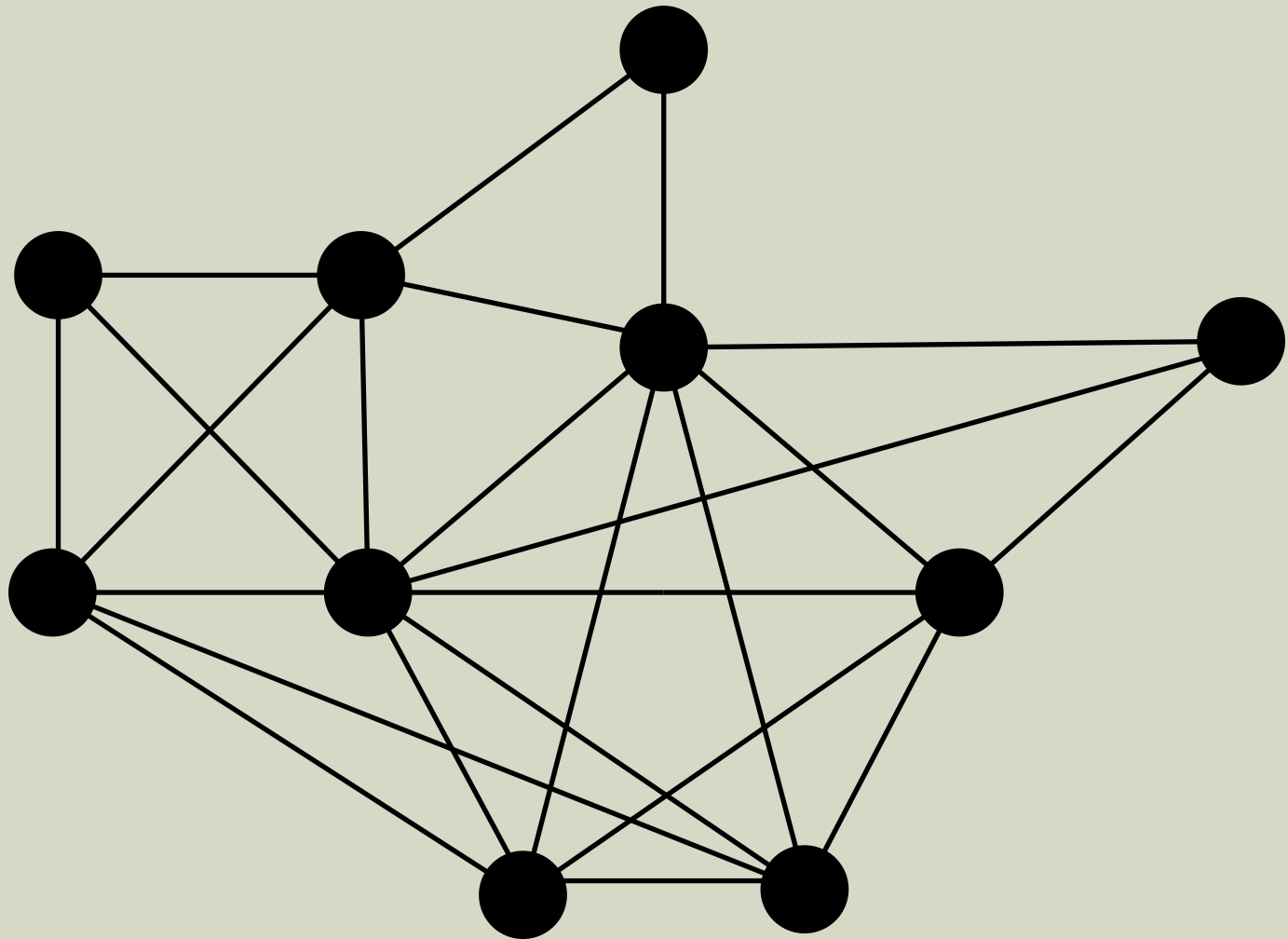


CLIQUE TO K-CLIQUE COMMUNITIES



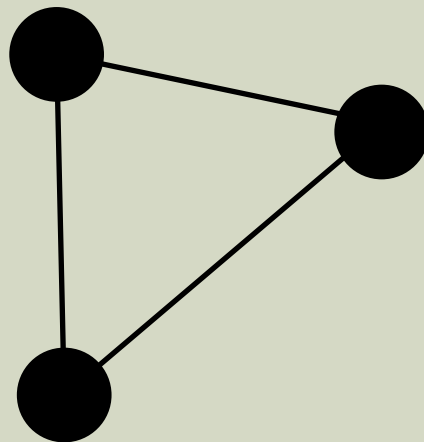
Clique 5: 3-clique

CLIQUEES TO K-CLIQUE COMMUNITIES



CLIQUE TO K-CLIQUE COMMUNITIES

Clique 6: 3-clique





A fast algorithm to do
form k -clique
communities

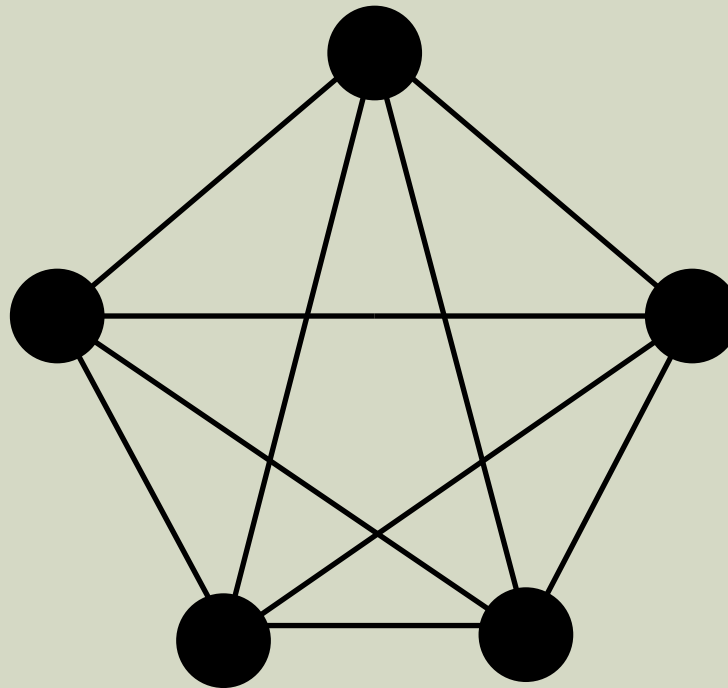
CLIQUE TO K-CLIQUE COMMUNITIES

	1	2	3	4	5	6
1	5					
2		4				
3			4			
4				4		
5					3	
6						3

CLIQUEES TO K-CLIQUE COMMUNITIES

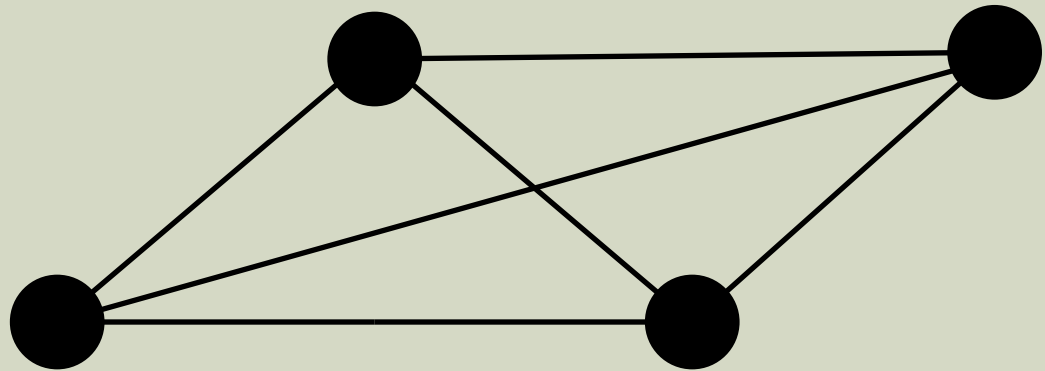
	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	3	2
6	2	2	2	1	2	3

CLIQUE TO K-CLIQUE COMMUNITIES



Clique 1: 5-clique

CLIQUE TO K-CLIQUE COMMUNITIES



Clique 2: 4-clique

CLIQUE TO K-CLIQUE COMMUNITIES

	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	3	2
6	2	2	2	1	2	3

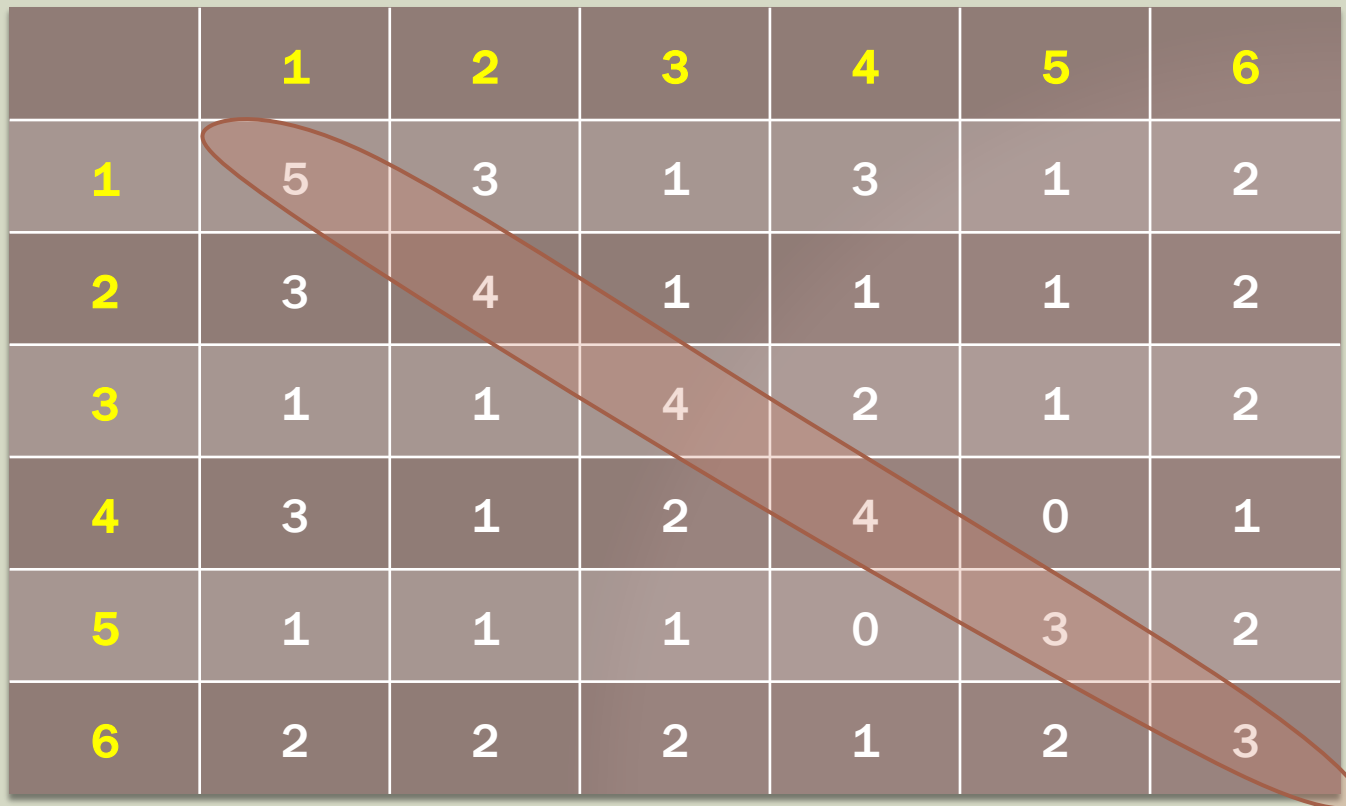
CLIQUEES TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	3	2
6	2	2	2	1	2	3

CLIQUE TO K-CLIQUE COMMUNITIES

$k=4$

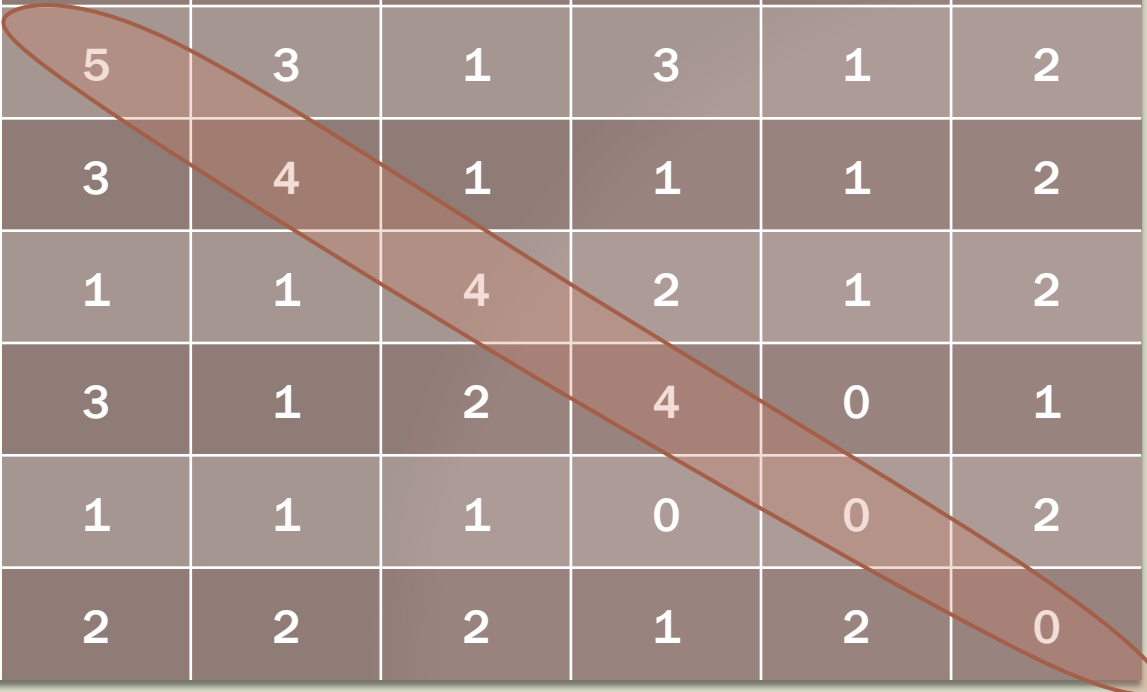


The image shows a 6x6 adjacency matrix for a graph with 6 nodes. The nodes are labeled 1 through 6 in yellow in the first row and first column. The matrix is symmetric. A red oval highlights a 4x4 submatrix in the top-left corner, representing a clique of size 4. The values in this submatrix are: (1,1)=5, (1,2)=3, (1,3)=1, (1,4)=3; (2,1)=3, (2,2)=4, (2,3)=1, (2,4)=1; (3,1)=1, (3,2)=1, (3,3)=4, (3,4)=2; (4,1)=3, (4,2)=1, (4,3)=2, (4,4)=4. The remaining cells in the matrix contain values: (1,5)=1, (1,6)=2; (2,5)=1, (2,6)=2; (3,5)=1, (3,6)=2; (4,5)=0, (4,6)=1; (5,5)=3, (5,6)=2; (6,5)=2, (6,6)=3.

	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	3	2
6	2	2	2	1	2	3

CLIQUEES TO K-CLIQUE COMMUNITIES

$k=4$



	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	0	2
6	2	2	2	1	2	0

Delete if less than k

CLIQUE TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	0	2
6	2	2	2	1	2	0

CLIQUE TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	5	3	1	3	1	2
2	3	4	1	1	1	2
3	1	1	4	2	1	2
4	3	1	2	4	0	1
5	1	1	1	0	0	2
6	2	2	2	1	2	0

CLIQUE TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	5	3	0	3	0	0
2	3	4	0	0	0	0
3	0	0	4	0	0	0
4	3	0	0	4	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Delete if less than $k-1$

CLIQUEES TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	5	3	0	3	0	0
2	3	4	0	0	0	0
3	0	0	4	0	0	0
4	3	0	0	4	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

CLIQUEES TO K-CLIQUE COMMUNITIES

$k=4$

	1	2	3	4	5	6
1	1	1	0	1	0	0
2	1	1	0	0	0	0
3	0	0	1	0	0	0
4	1	0	0	1	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Change all non-zeros to 1

CLIQUE TO K-CLIQUE COMMUNITIES

$k=4$

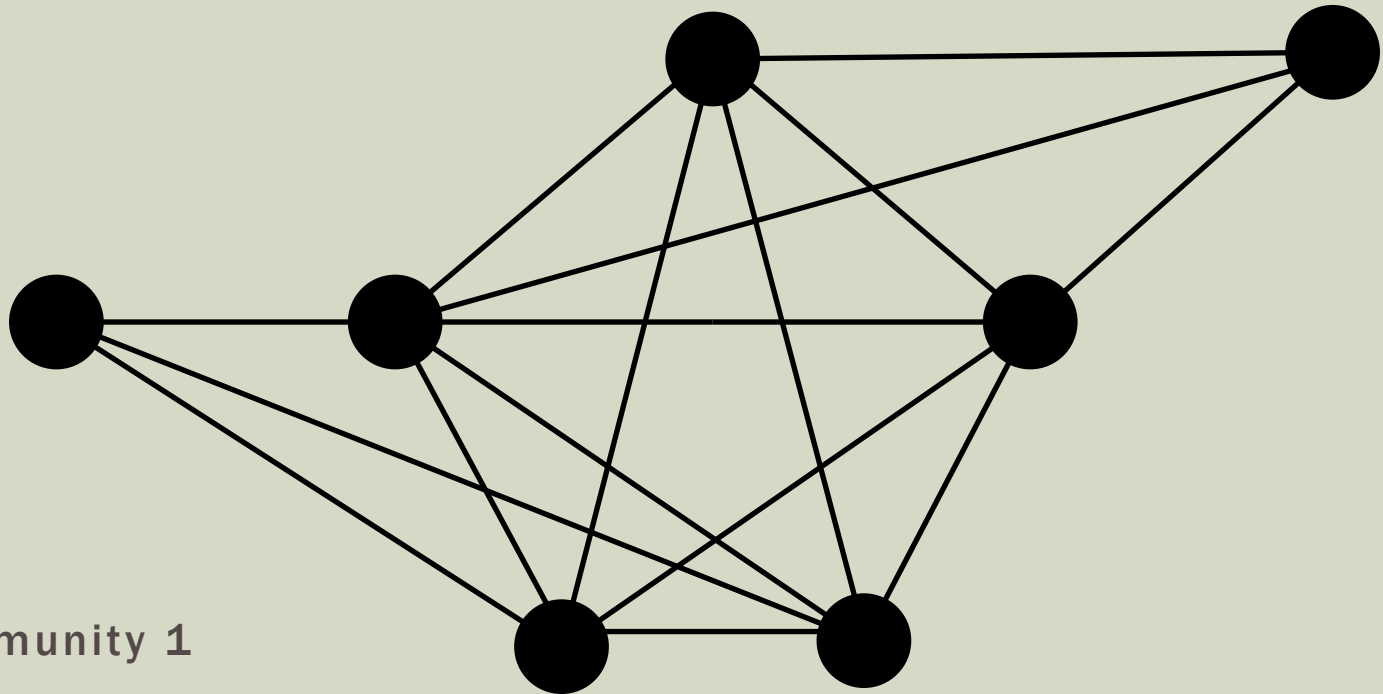
	1	2	3	4	5	6
1	1	1	0	1	0	0
2	1	1	0	0	0	0
3	0	0	1	0	0	0
4	1	0	0	1	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0

Clique-clique overlap matrix

CLIQUEES TO K-CLIQUE COMMUNITIES

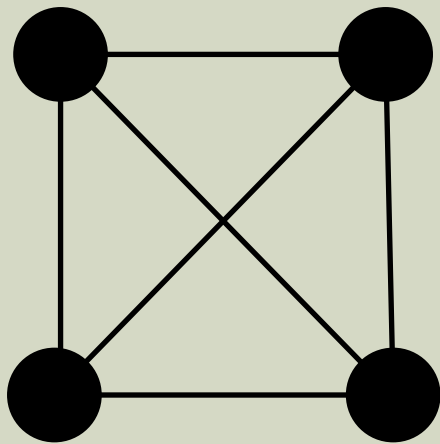
$k=4$

Community 1



CLIQUEES TO K-CLIQUE COMMUNITIES

$k=4$



Community 2

ANALYSIS

- Believed to be non-polynomial
- No closed formula can be given
- However, claimed to be efficient on real systems

CONCLUSION

- Widely used algorithm for detecting overlapping communities
- However:
 - Fail to give meaningful covers for graph with few cliques
 - With too many cliques, might give a trivial community structure
 - Left out vertices?
 - Subgraphs containing many cliques == community?
 - What value of k to choose to give a meaningful structure?

REFERENCES

- Palla et al. – Uncovering the overlapping community structure of complex networks in nature and society
- Santo Fortunato - Community detection in graphs

Wanna use Clique Percolation Method?

Just google: “cfinder”

[CFinder] Clusters and Communities: Overlapping dense groups in networks - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://angel.elte.hu/cfinder/

Most Visited Getting Started Latest Headlines Share on Facebook http://www.hardrock... chart (PNG Image, 50... http://128.103.191.2...

Community structure in social ... CA Hidalgo Publications of Gergely Palla community.pdf (application/p... gergely palla ppt - Google Sea... cesar hidalgo ppt - Google Sea... [CFinder] Clusters and Co... [CFinder] Clusters and Comm...

CURRENT VERSION:
CFinder 2.0.1
Jan 02, 2009
Requires Java >>

Clusters & Communities

overlapping dense groups in networks

Download CFinder Manual Network Data Publications WebCFinder

Home

Download

Software
Network Data

Publications

Support

FAQ
Manual
Contact us

People

Links

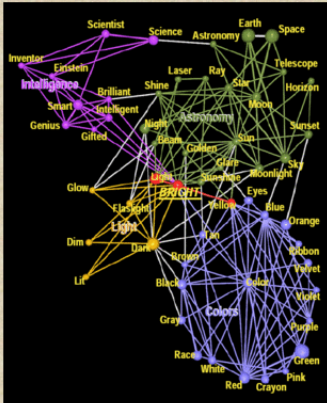
CFinder is a free software for finding and visualizing overlapping dense groups of nodes in networks, based on the Clique Percolation Method (CPM) of Palla et. al., *Nature* **435**, 814-818 (2005).

NEW CFinder has been recently applied to the quantitative description of the evolution of social groups: Palla et. al., *Nature* **446**, 664 (2007).

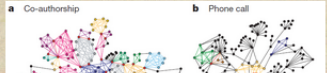
CFinder offers a fast and efficient method for clustering data represented by large graphs, such as genetic or social networks and microarray data. CFinder is also very efficient for locating the cliques of large sparse graphs.

Download: [Software](#) | [Manual](#) | [Publications](#)

A cluster -- also called a community or module -- in a network is a group of nodes more densely connected to each other than to nodes outside the group. In real networks clusters often overlap.



The **overlapping network modules** of the word "bright" in a word association network represent the different meanings of this word. From Palla et. al., *Nature* **435**, 814-818 (2005).



Thank you!