Stad: Stateful Diffusion for Linear Time Community Detection.
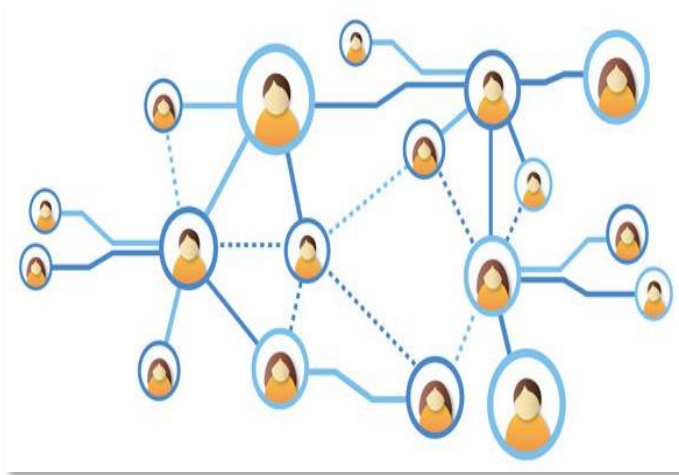## ICDCS'18
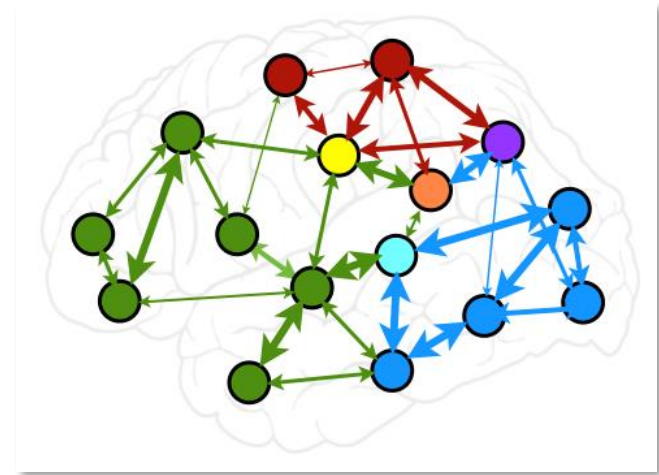
Amira Soliman, Fatemeh Rahimian, and Sarunas Girdzijauskas

DISCAN'18, Stockholm, Sweden.
Sep 26 – 27, 2018.
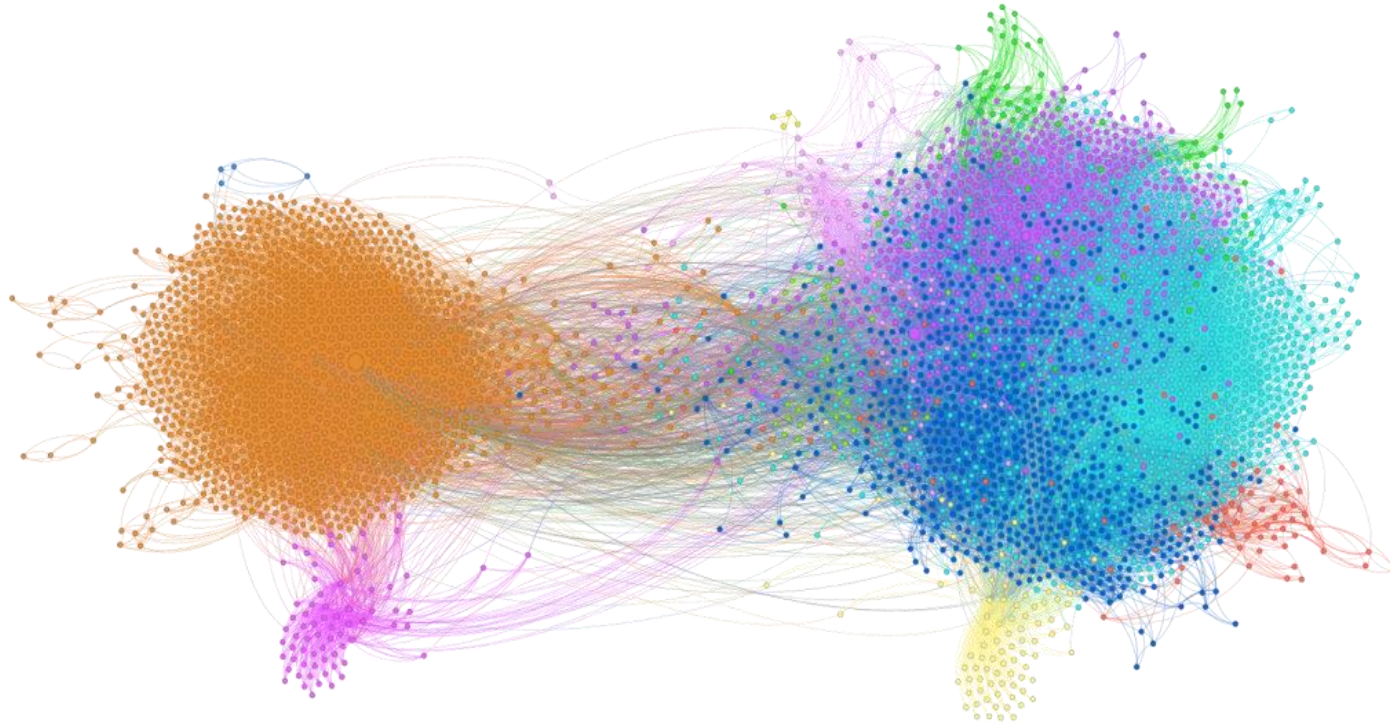
Network Analysis

Social networks

Biological networks

Community Analysis

- **Community represents a group of nodes that are densely connected to each other and sparsely connected to the rest of the graph.**

Community Analysis

- **Community represents a group of nodes that are densely connected to each other and sparsely connected to the rest of the graph.**
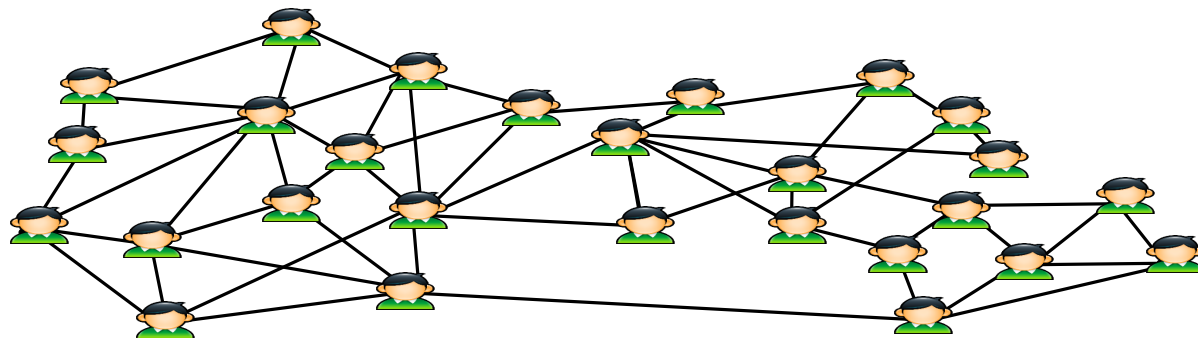- **Communities vary in size and internal cohesion.**
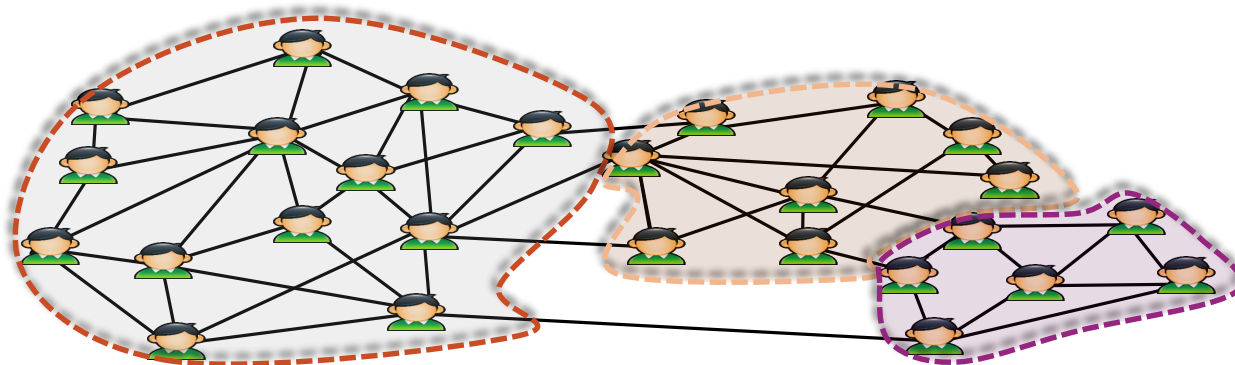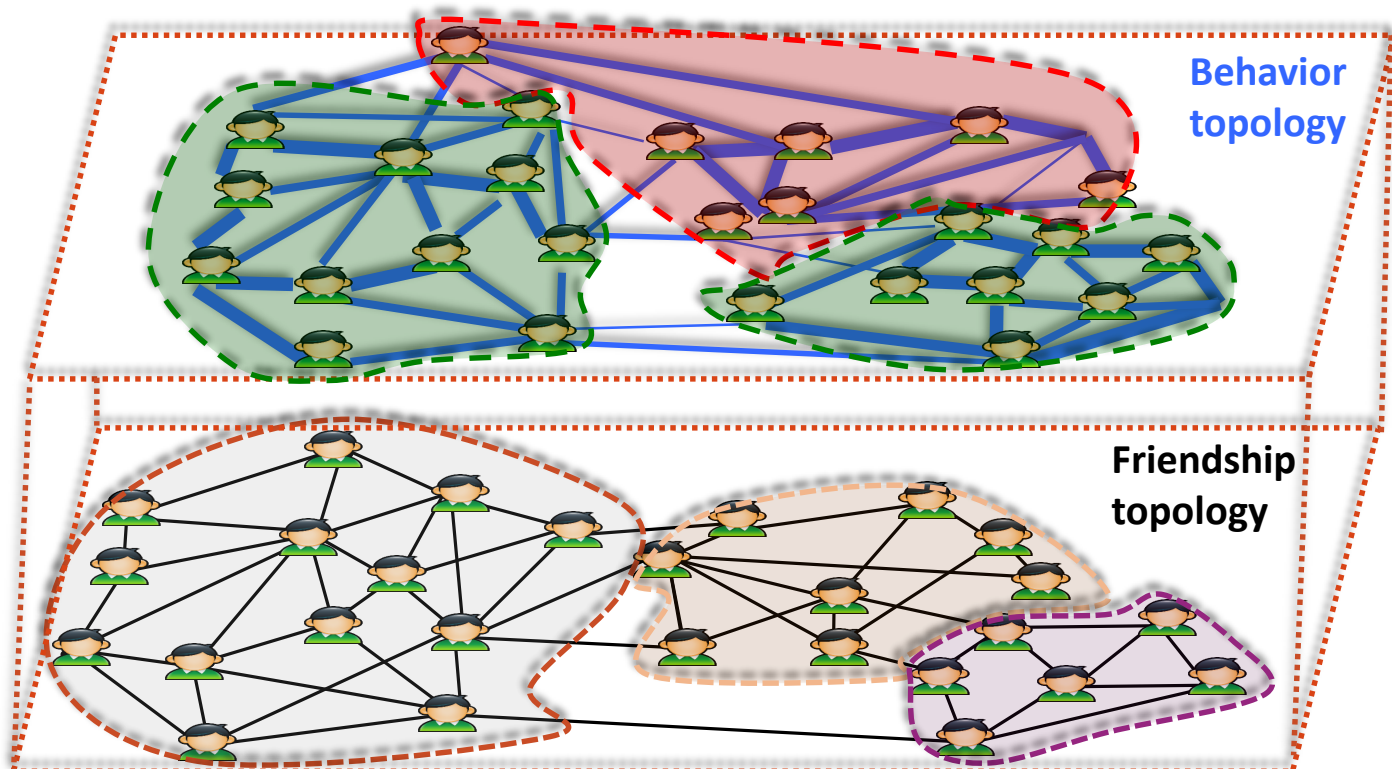
Community-aware Learning

Integrating community structure with analytic tasks enhances the data analytic insights and improves the results (e.g., recommender systems and spam detection).

Community-aware Learning

Integrating community structure with analytic tasks enhances the data analytic insights and improves the results (e.g., recommender systems and spam detection).
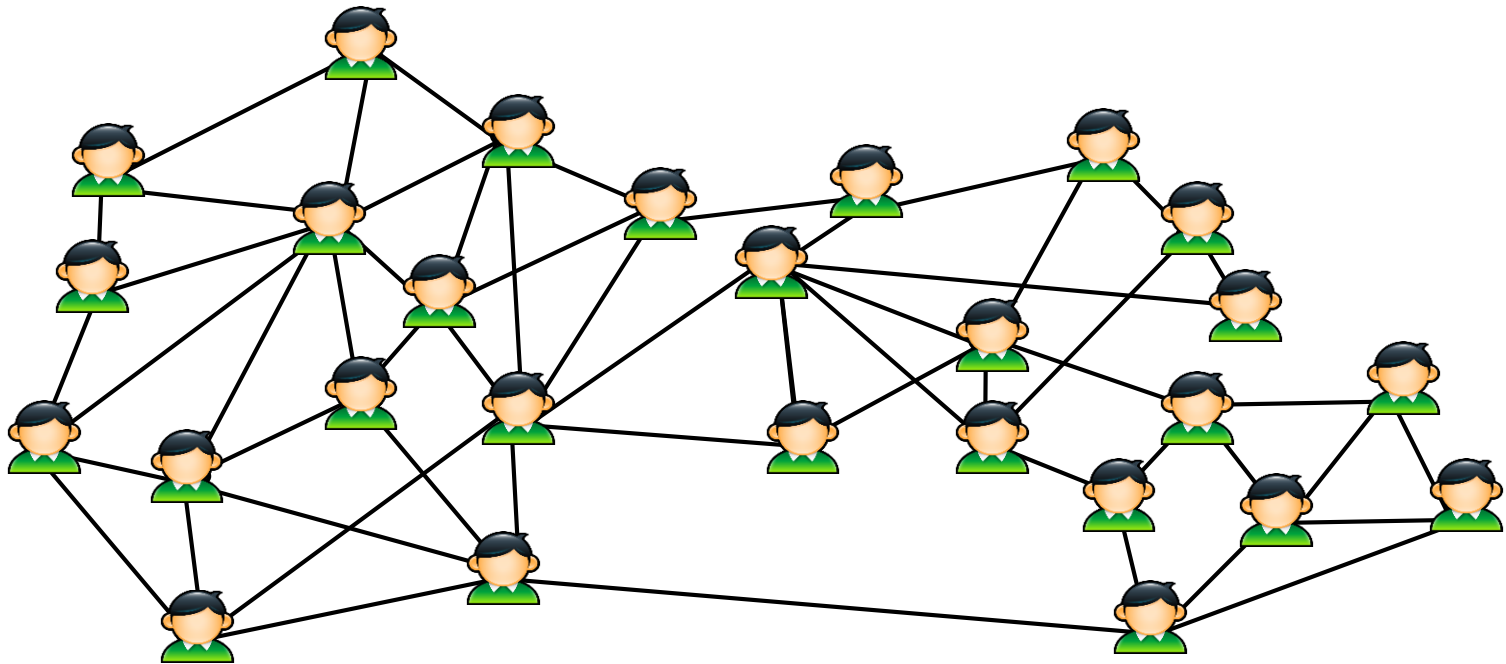
Community-aware Learning

Integrating community structure with analytic tasks enhances the data analytic insights and improves the results (e.g., recommender systems and spam detection).
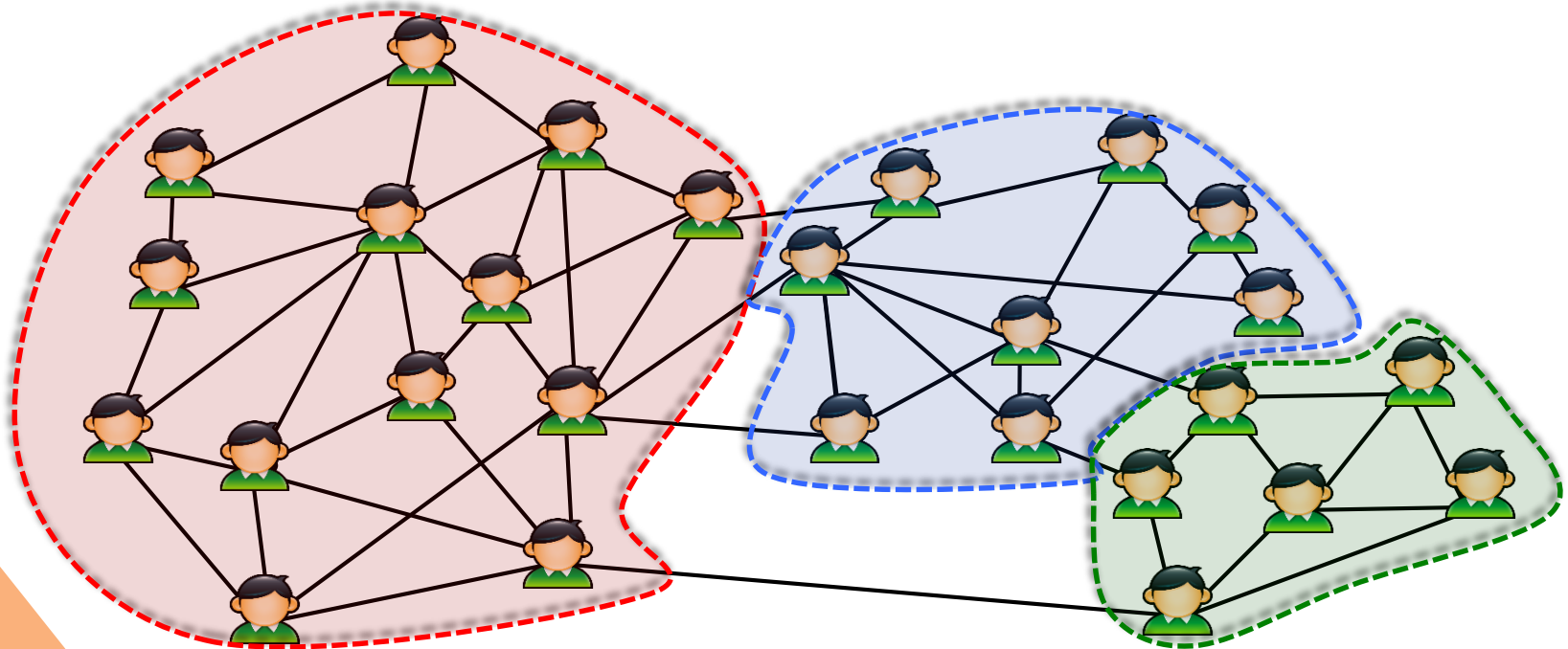
Community Detection

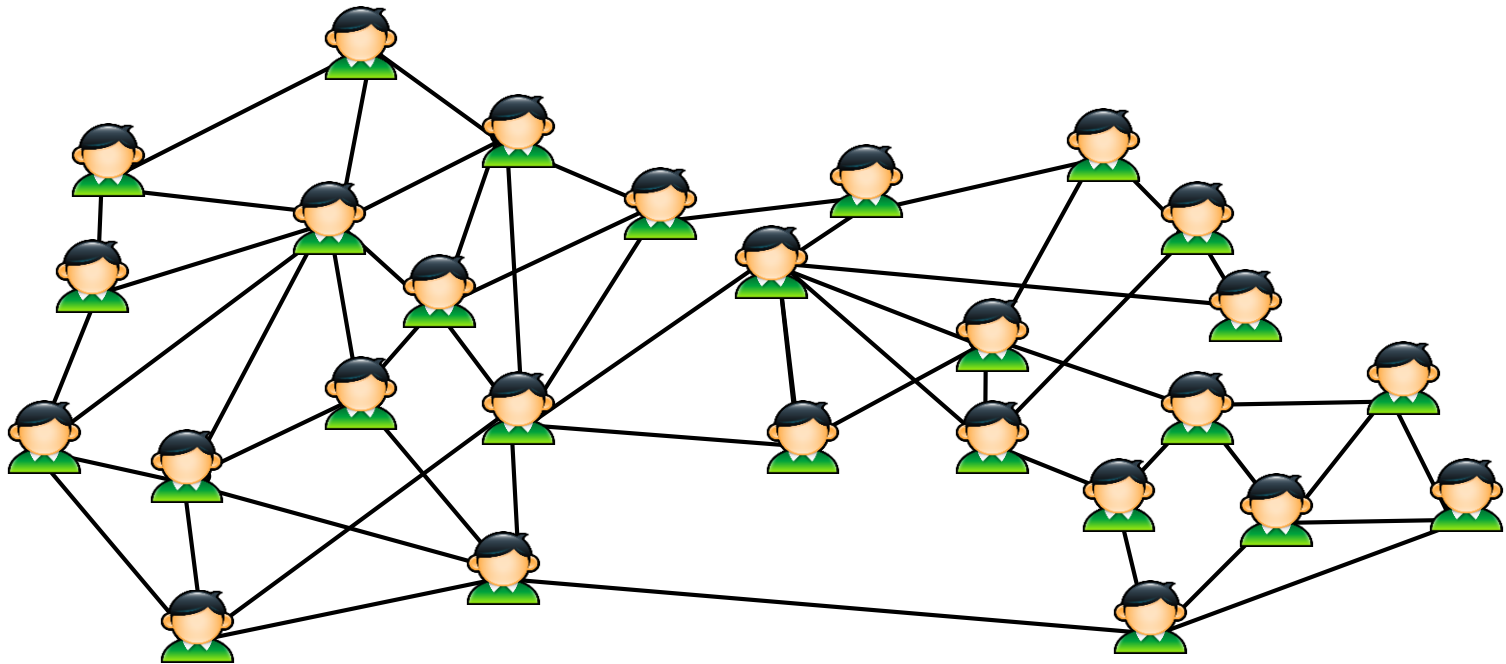**A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.**

Community Detection

A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.
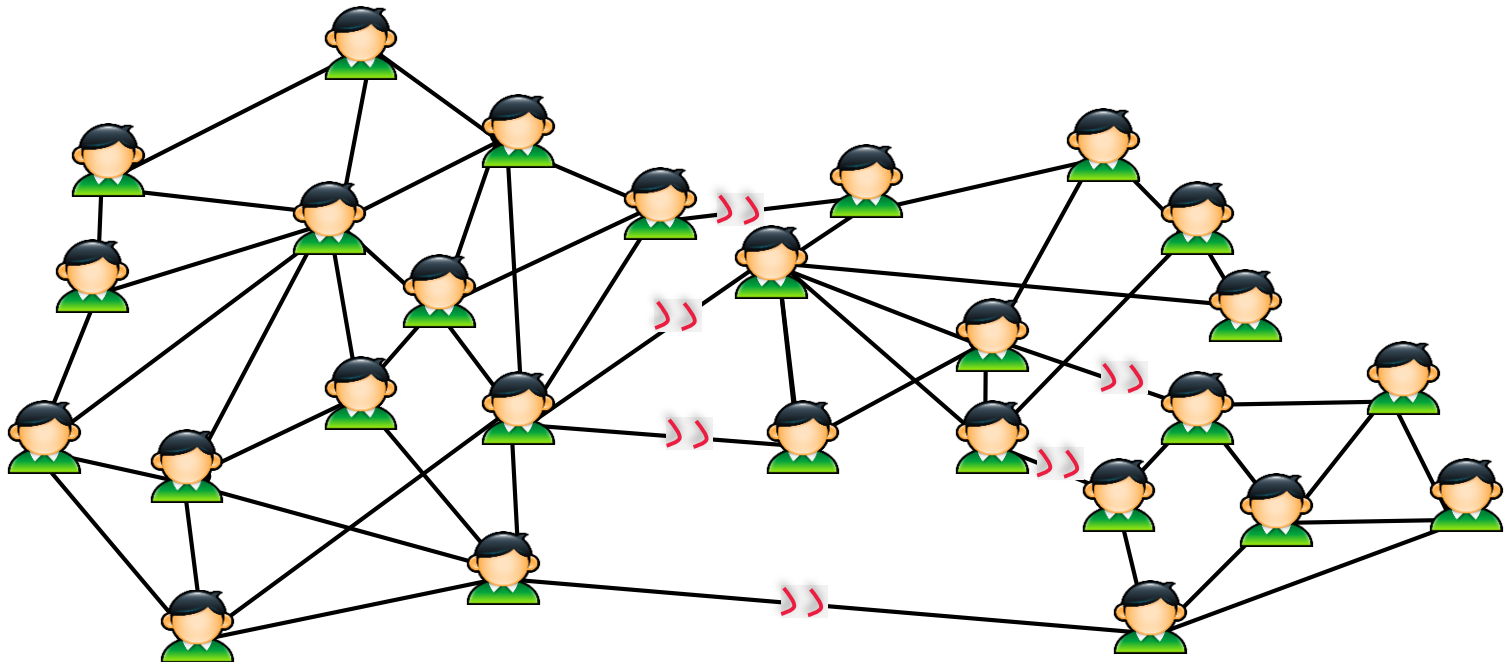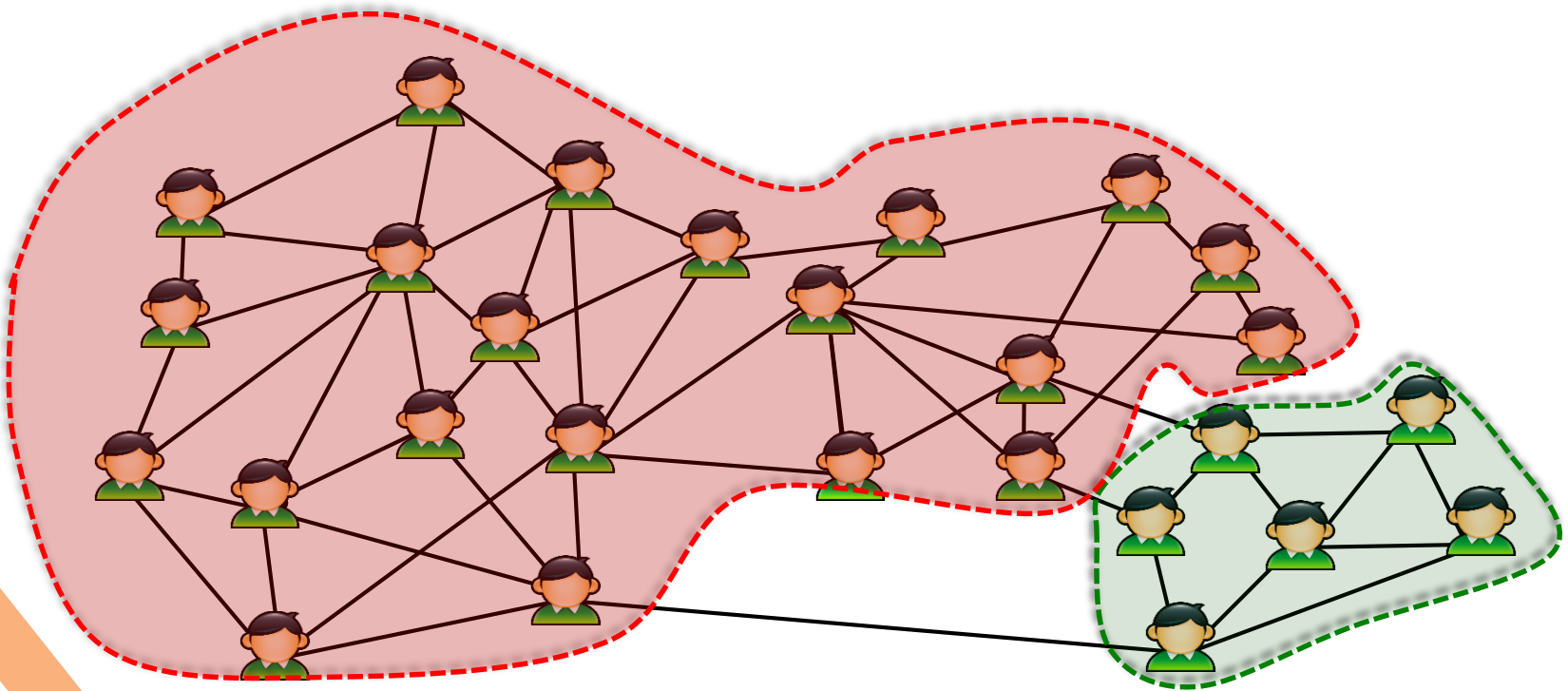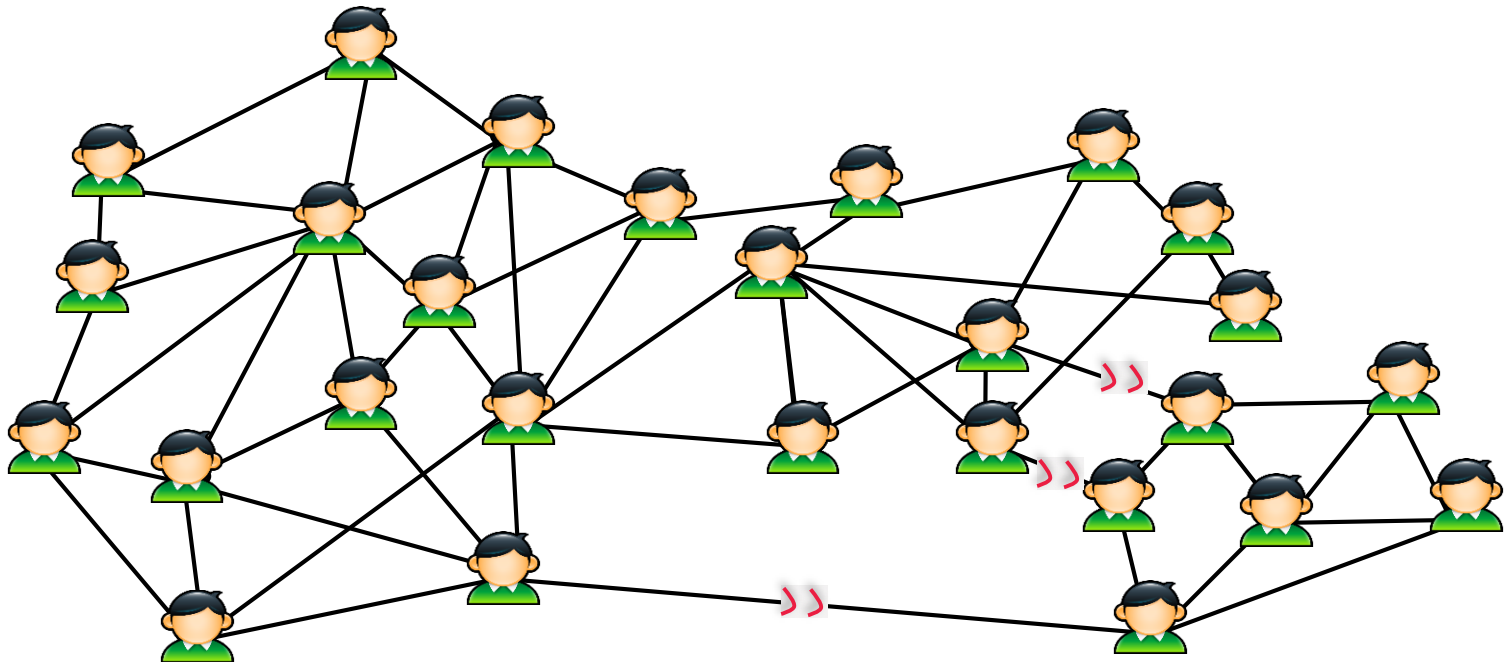
Community Detection

A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.

Community Detection

A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.
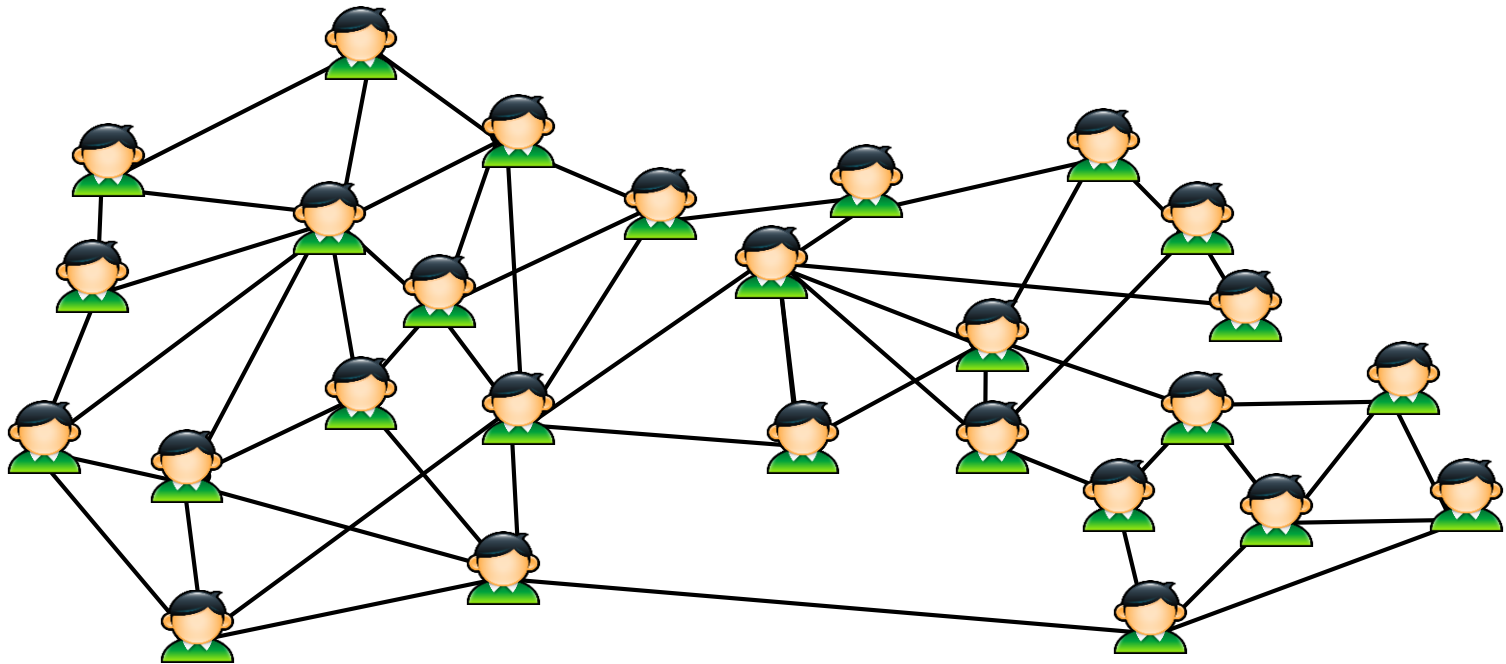
Community Detection

A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.

Community Detection

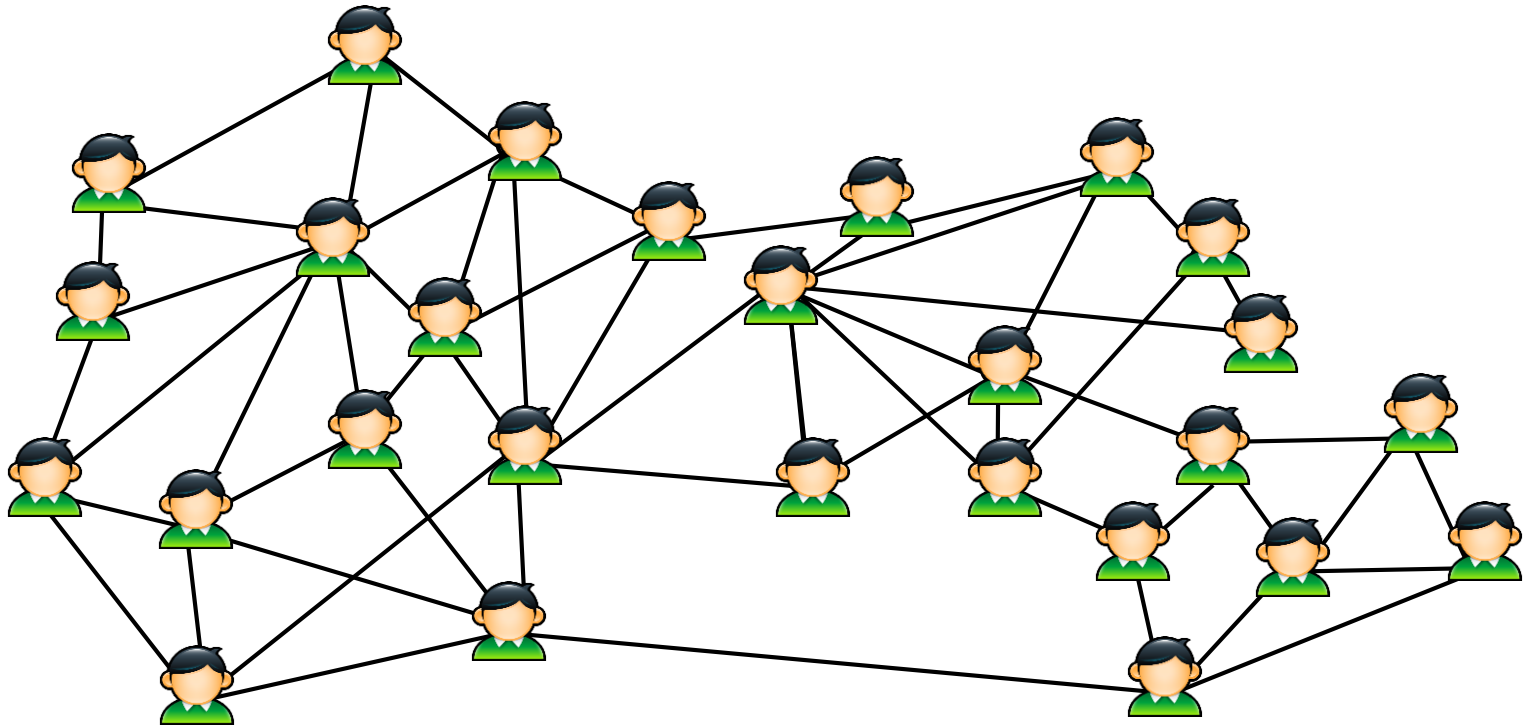**A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.**

Community Detection

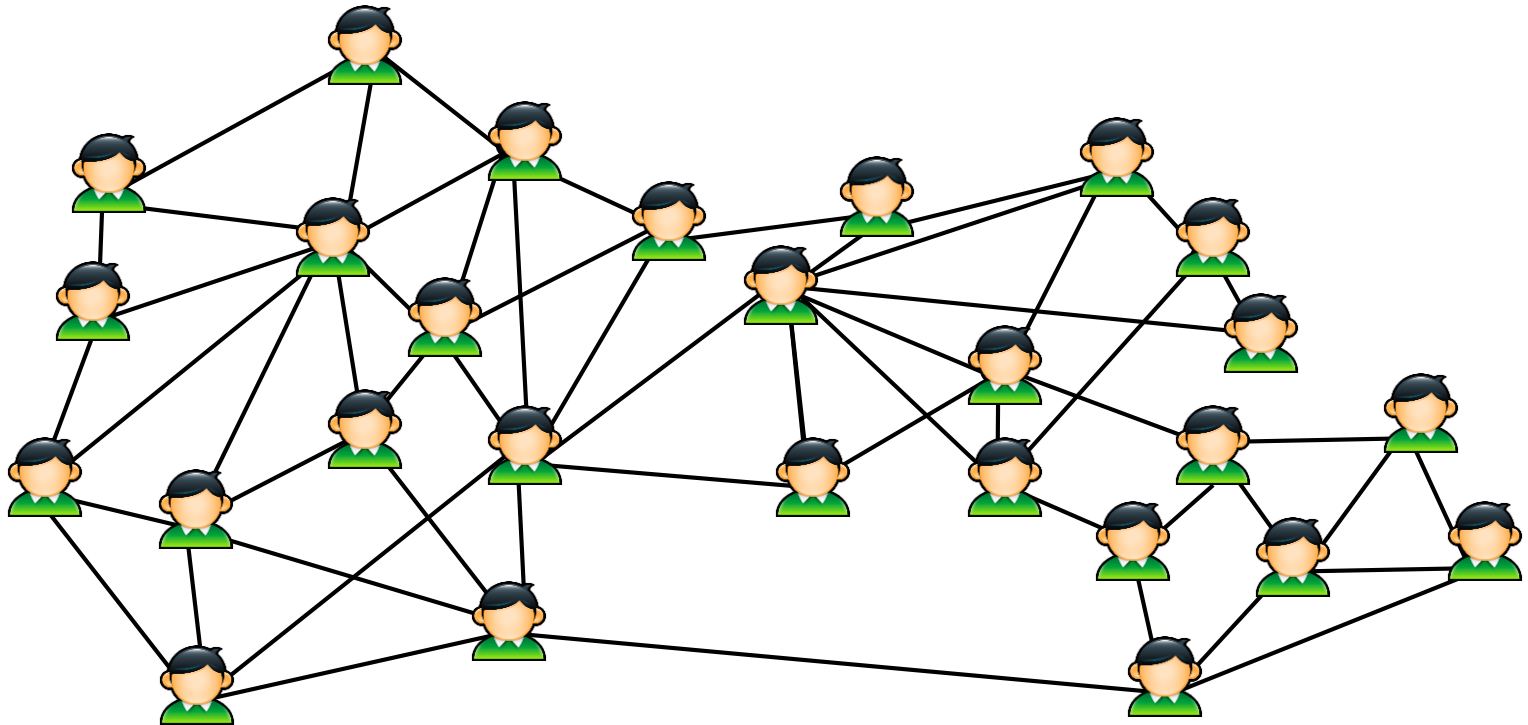**A community corresponds to a group of nodes with more intra-cluster than inter-clusters edges.**

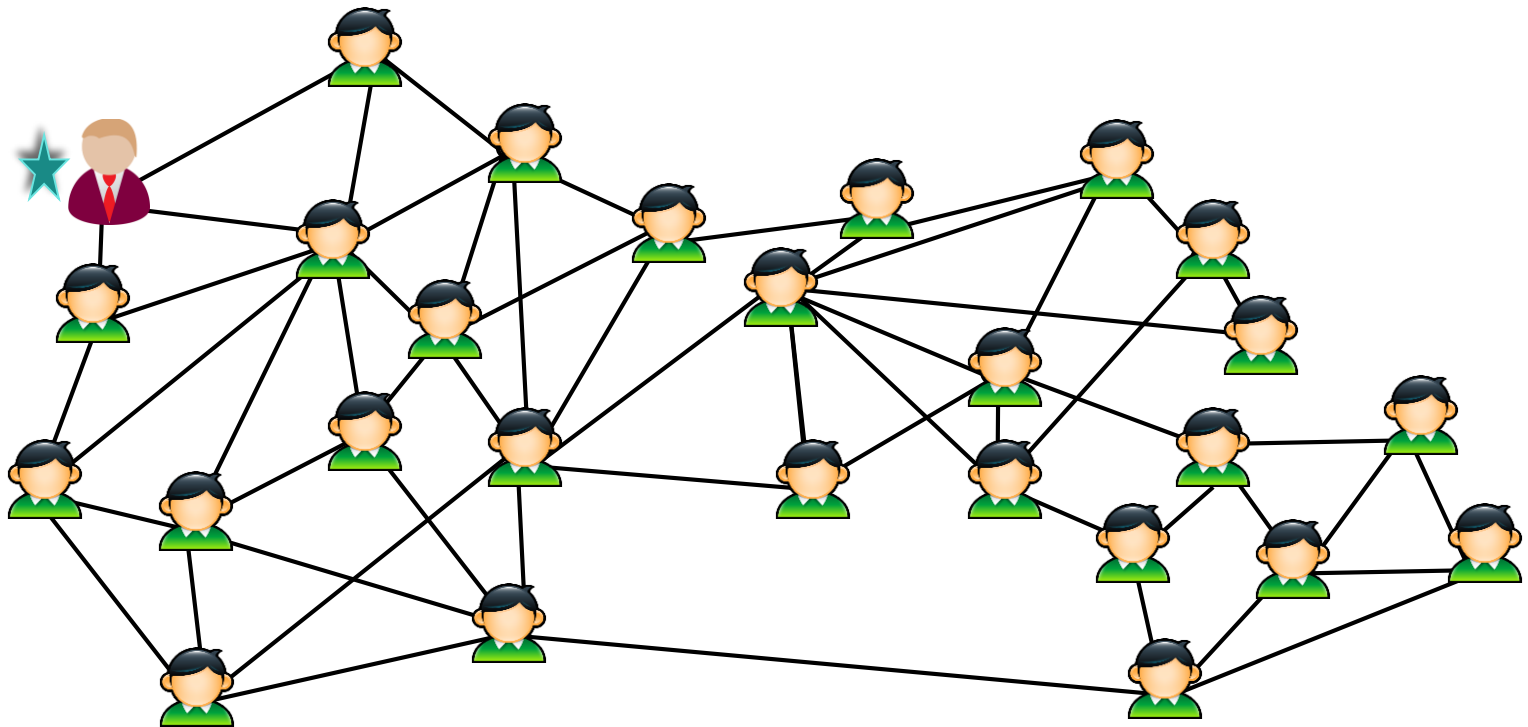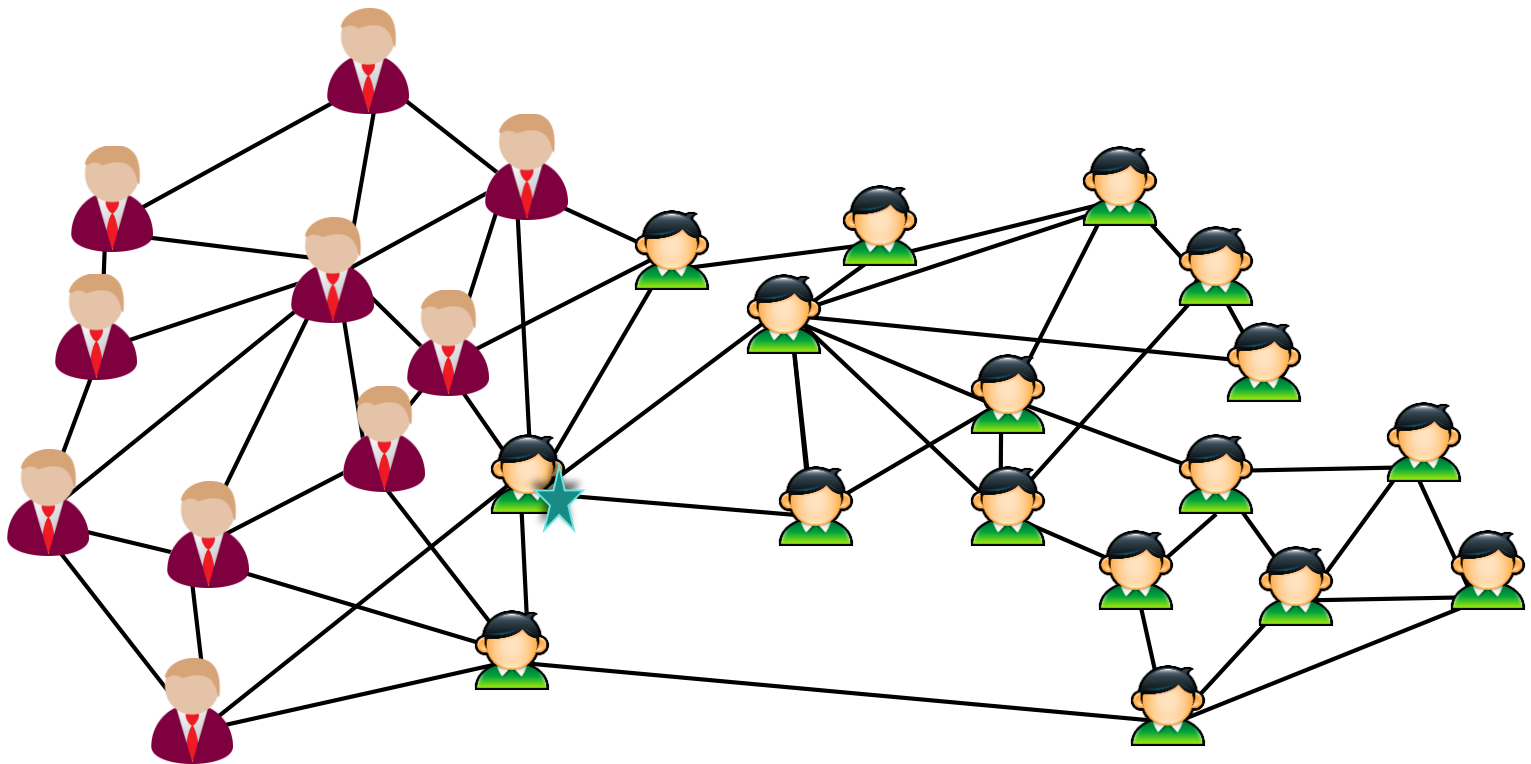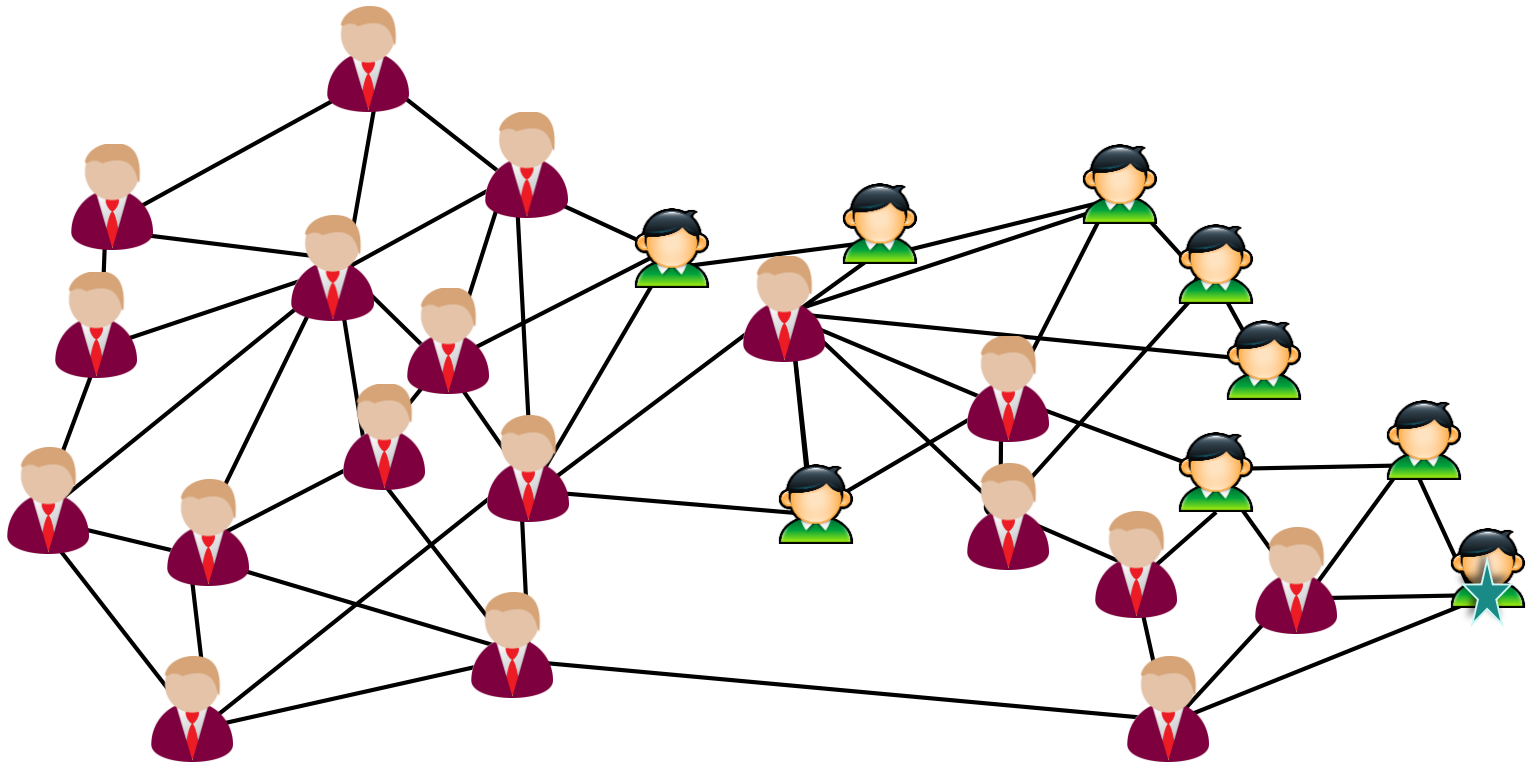# Localized-methods for Community Detection

Localized-methods for Community Detection

Random Walks (RWs):

Localized-methods for Community Detection

Random Walks (RWs):
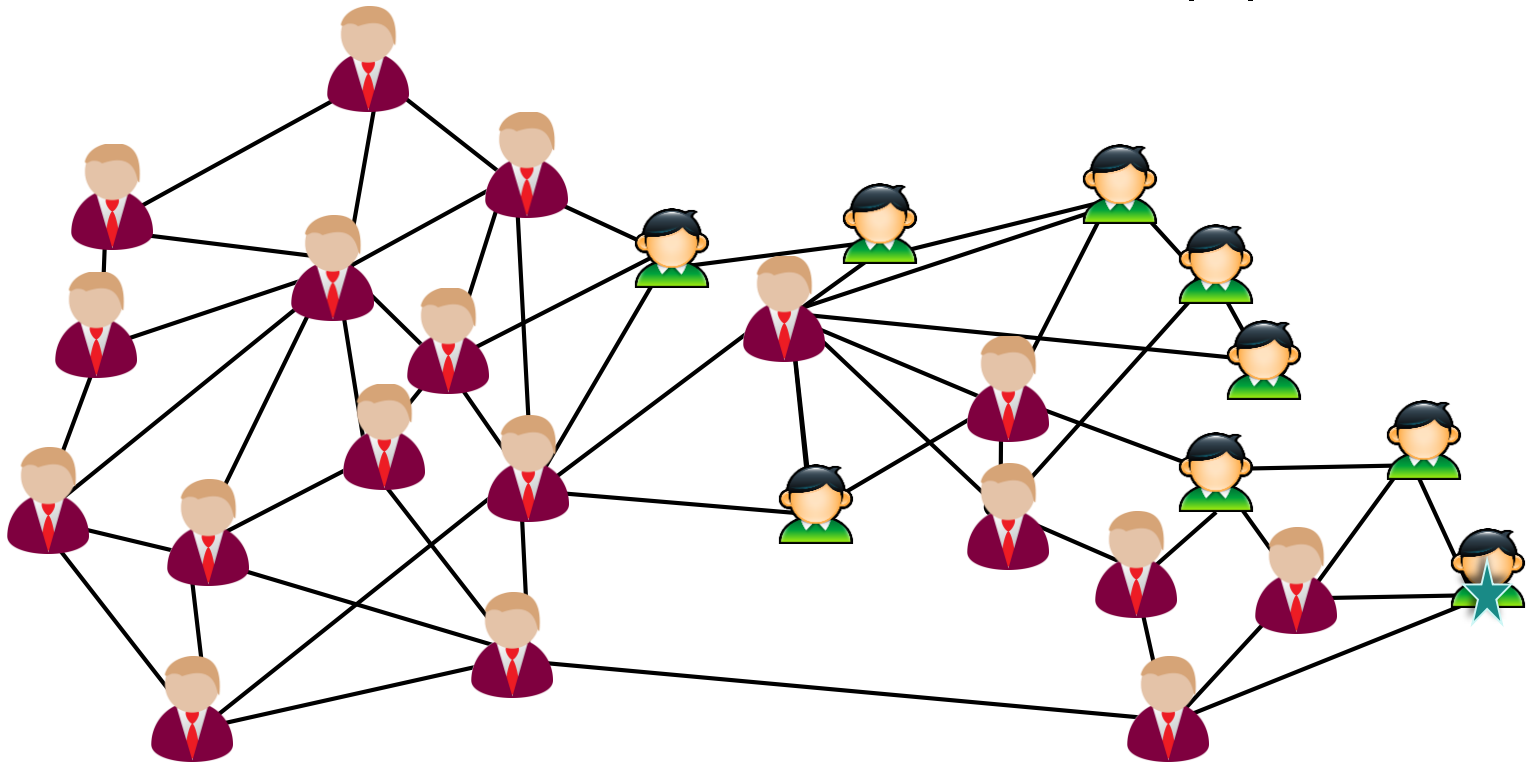
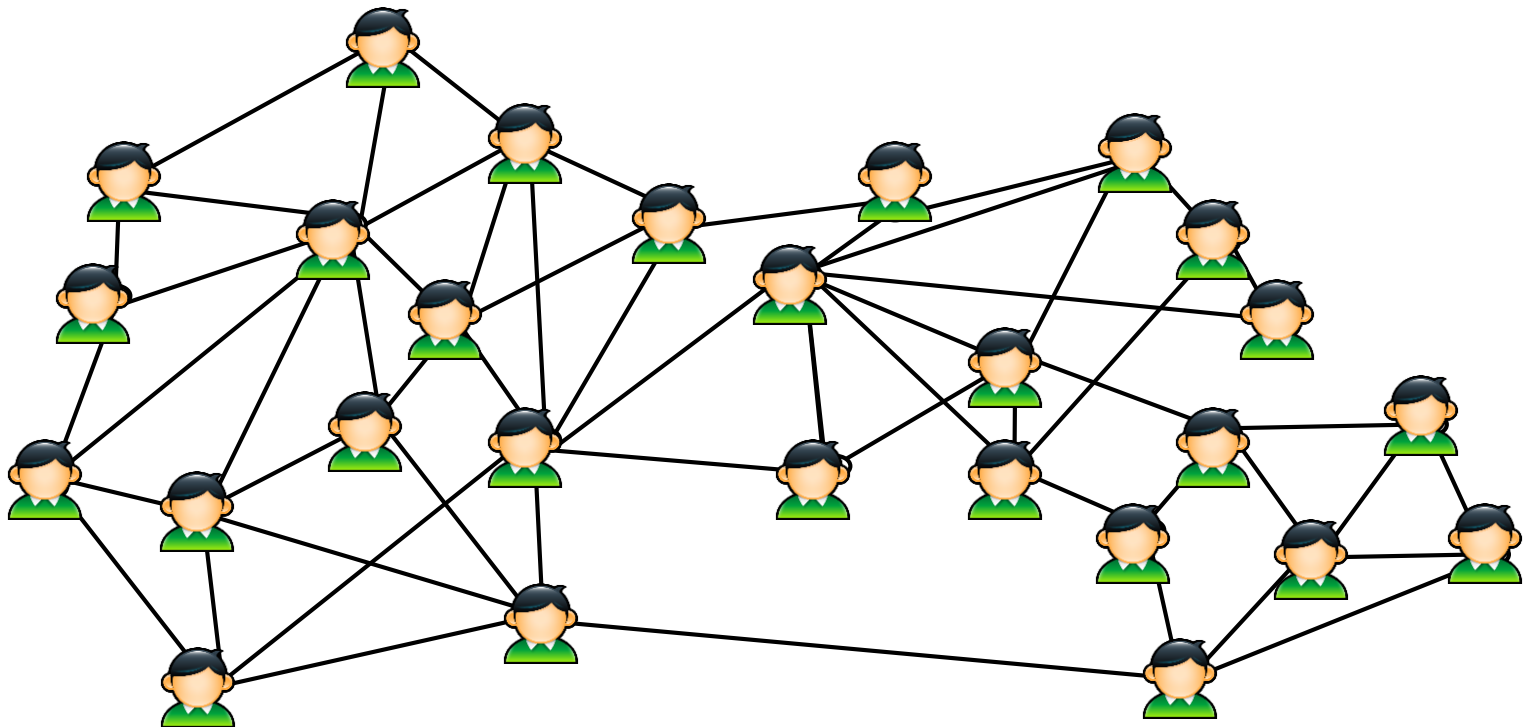# Localized-methods for Community Detection

Random Walks (RWs):

Localized-methods for Community Detection

Random Walks (RWs):

Localized-methods for Community Detection

Random Walks (RWs):
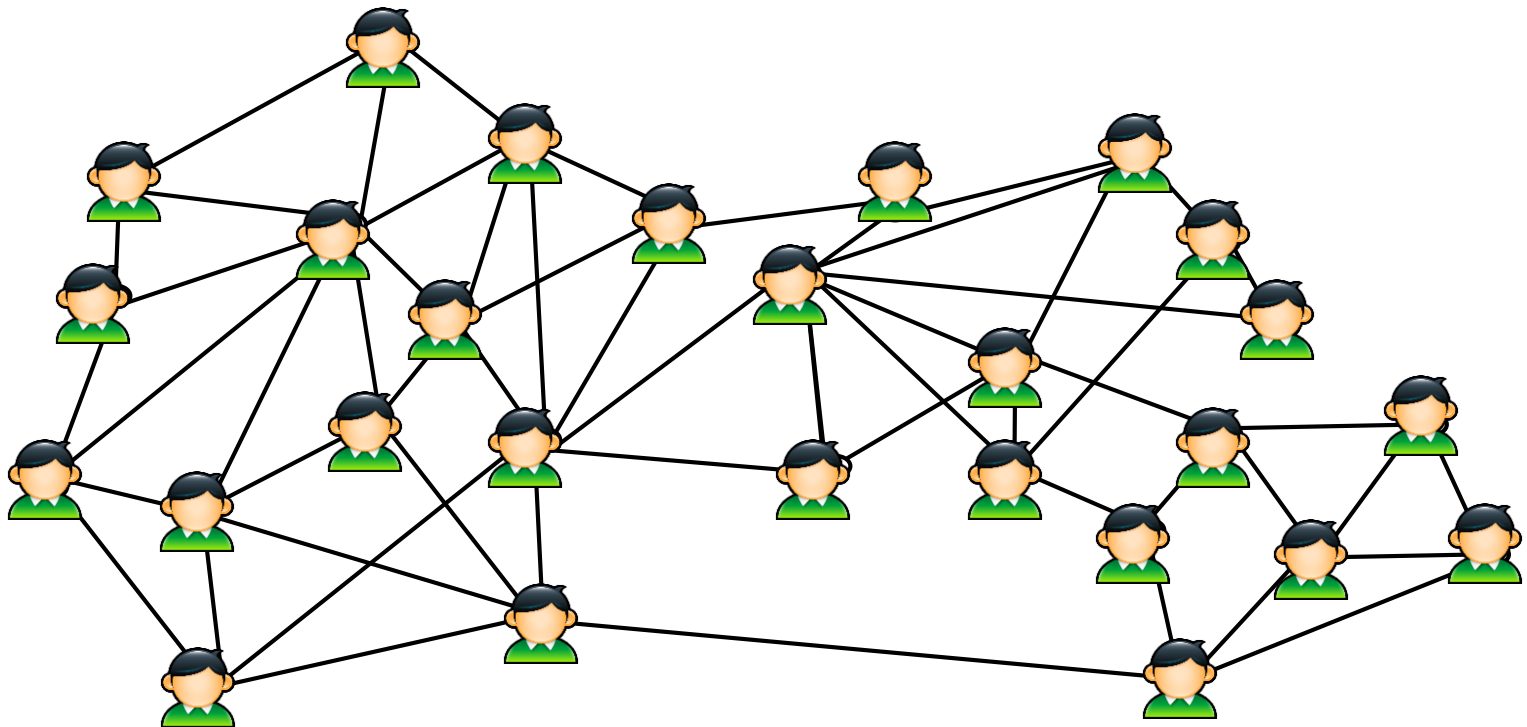
$$\pi\,(v) = \frac{d_v}{2|E|}$$

Localized-methods for Community Detection

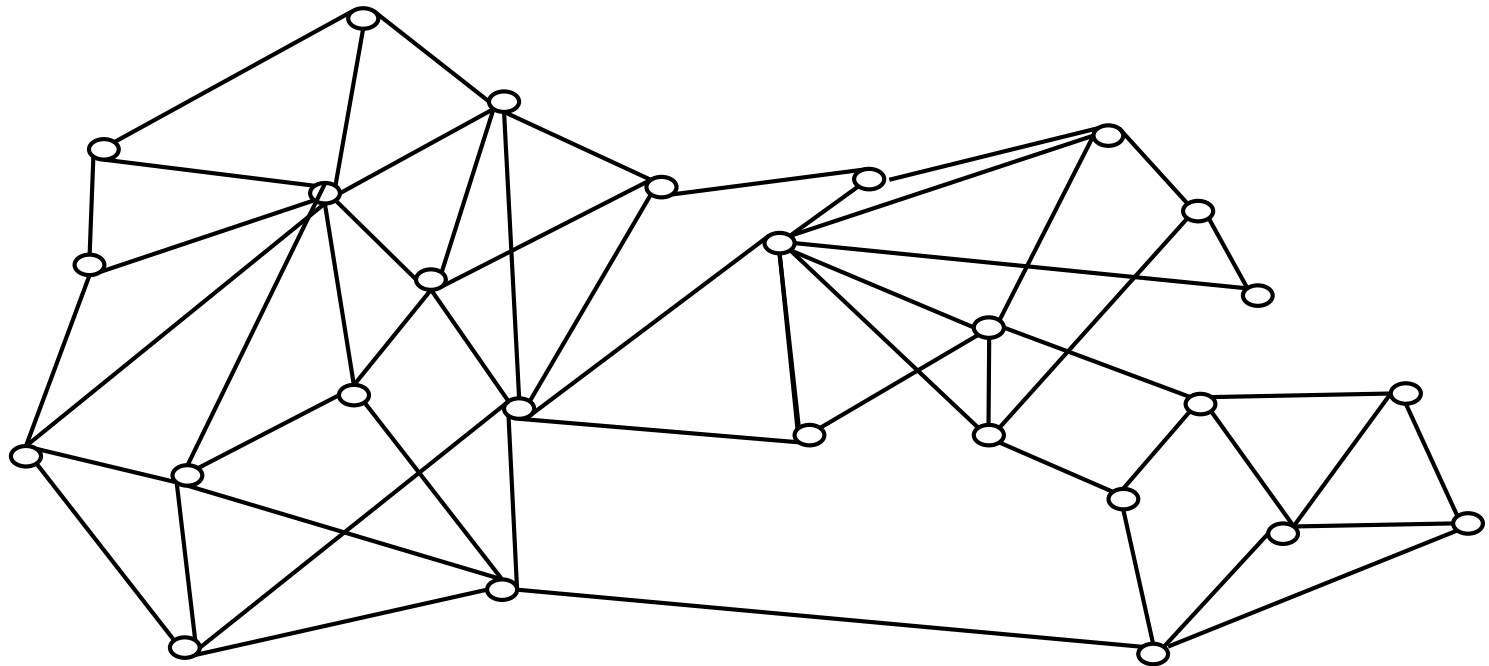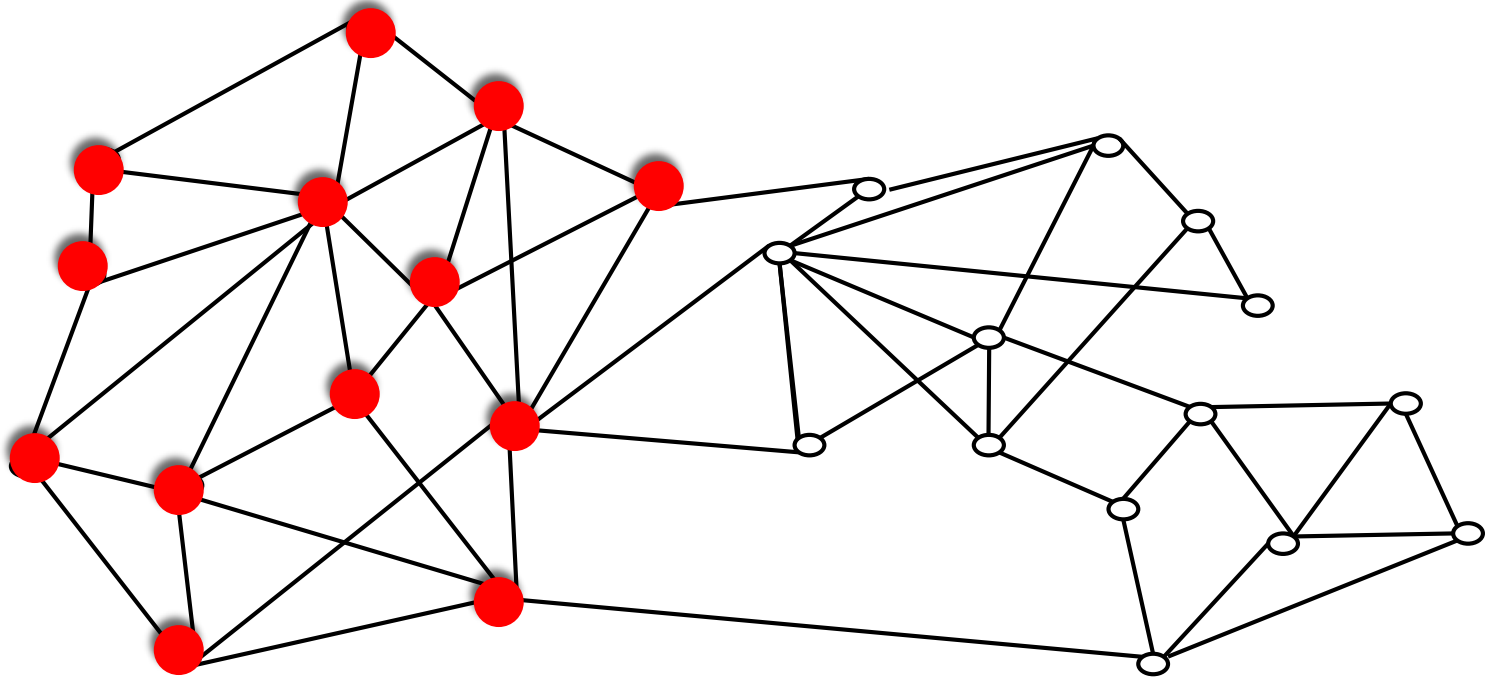# Localized-methods for Community Detection

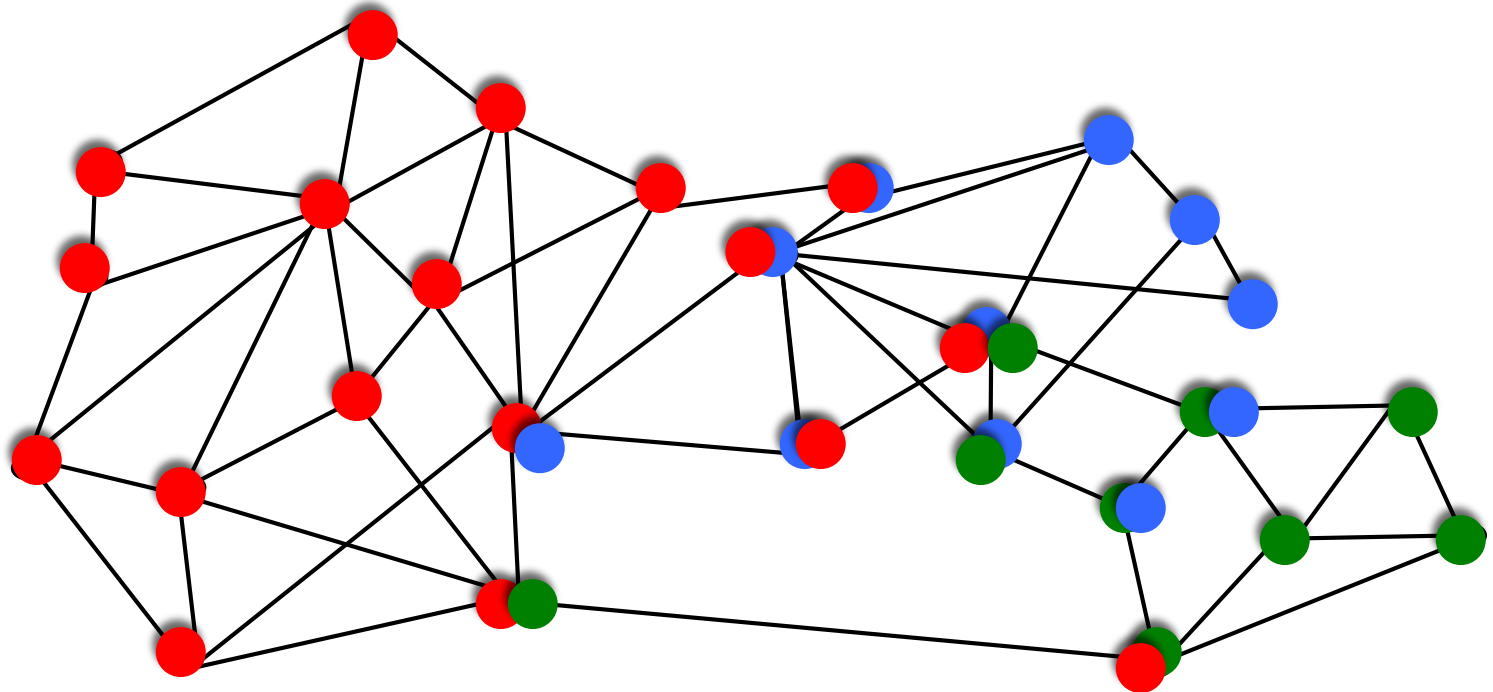Diffusion:

Localized-methods for Community Detection

Diffusion:

Localized-methods for Community Detection

Diffusion:

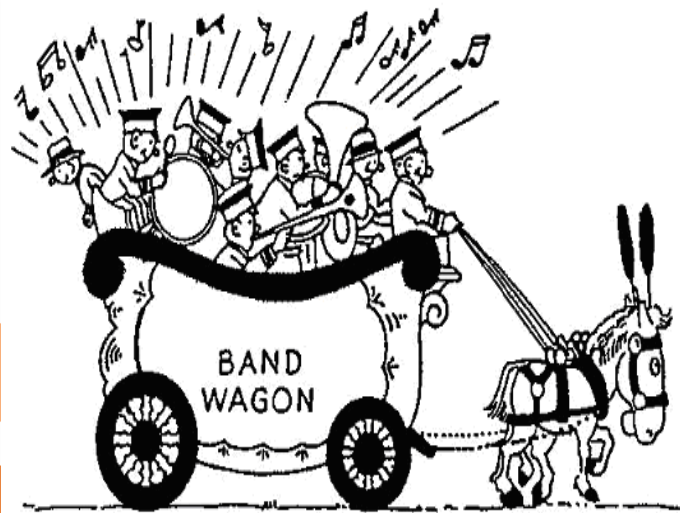# Localized-methods for Community Detection

Diffusion:

Bandwagon Effect?

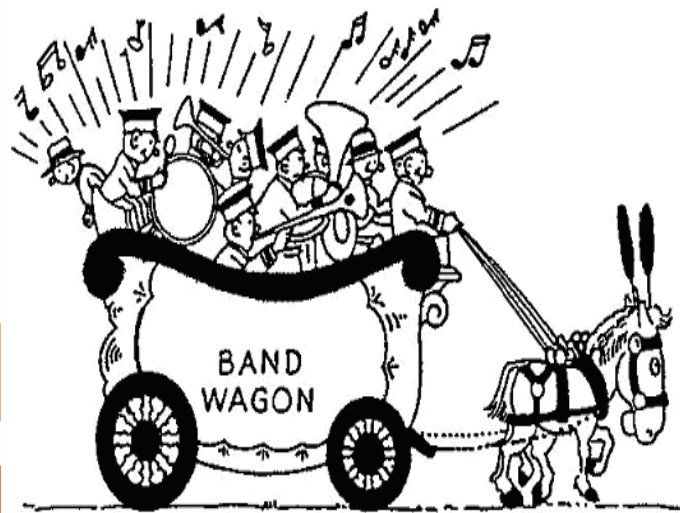# Bandwagon Effect?

Bandwagon Effect?

- **It is a cognitive bias that refers to the tendency to follow other people just because they are the majority.**

Bandwagon Effect?

- **It is a cognitive bias that refers to the tendency to follow other people just because they are the majority.**

- Heterogenous Community Size!
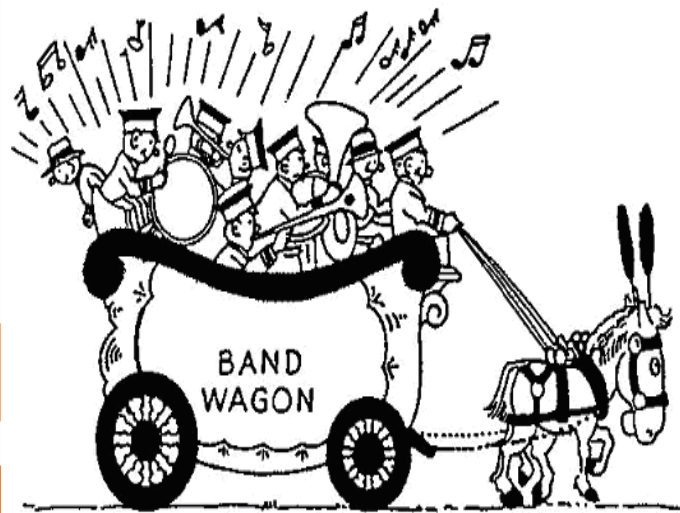
Bandwagon Effect?

- **It is a cognitive bias that refers to the tendency to follow other people just because they are the majority.**
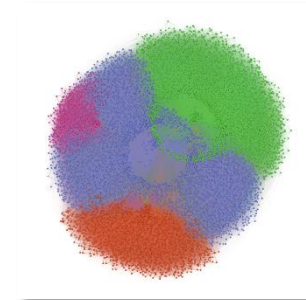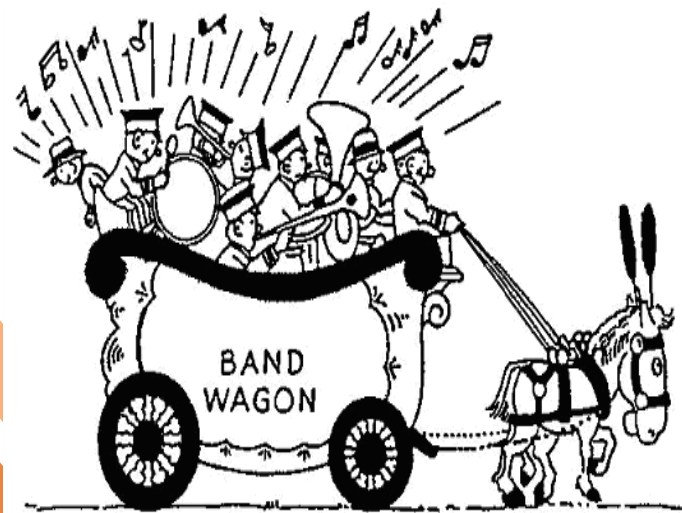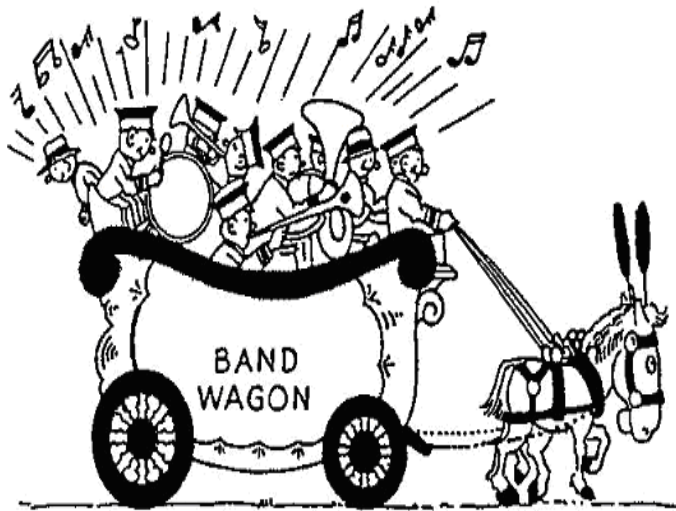

- Heterogenous Community Size!

Ground-truth

Bandwagon Effect?

- **It is a cognitive bias that refers to the tendency to follow other people just because they are the majority.**

- Heterogenous Community Size!

Ground-truth

Resolution limit

Bandwagon Effect?

- **It is a cognitive bias that refers to the tendency to follow other people just because they are the majority.**
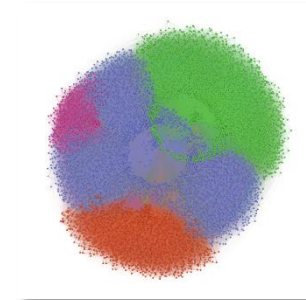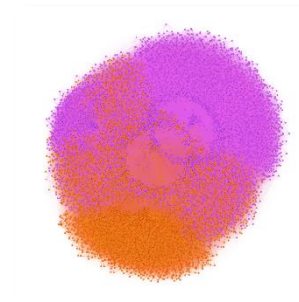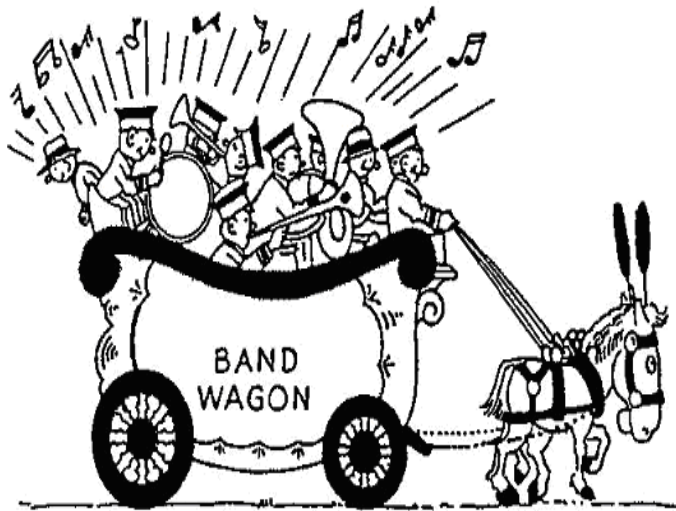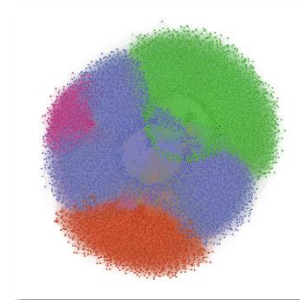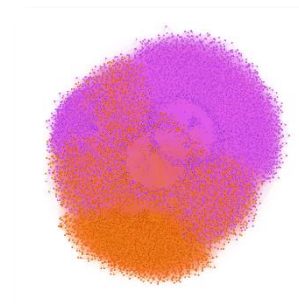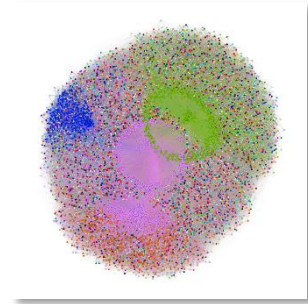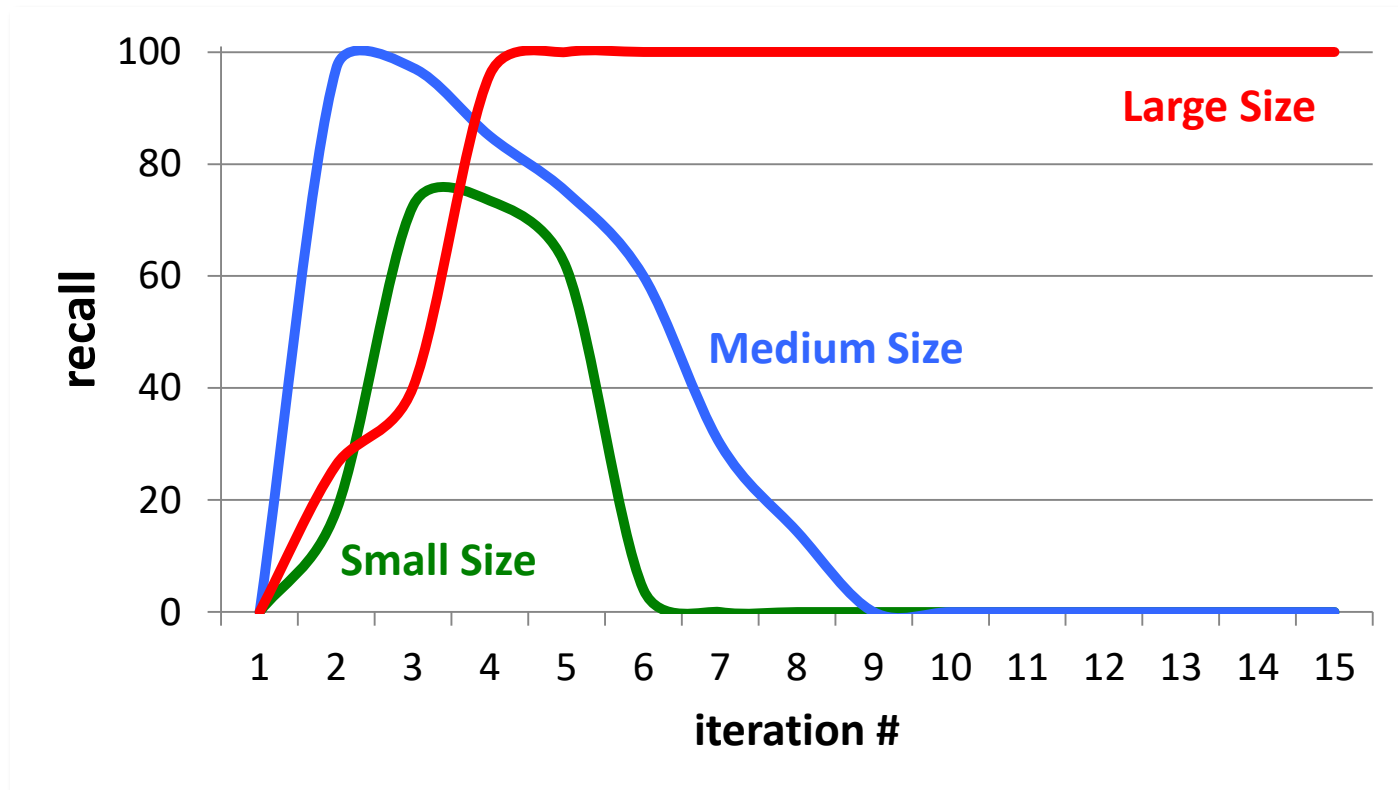

- Heterogenous Community Size!

Ground-truth

Resolution limit

Field-of-view limit

Heterogenous Community Size

Random Walks and Diffusion Performance!

Adaptive Diffusion

Adaptive Diffusion

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.

Colors	Units
Red	
Green	
Blue	

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.

Colors	Units
Red	
Green	
Blue	

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.

Colors	Units
Red	■■■■■■■
Green	■
Blue	■■

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.
2. Decide membership by selecting the dominant color.

Colors	Units
Red	■■■■■■■
Green	■
Blue	■■

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.
2. Decide membership by selecting the dominant color.

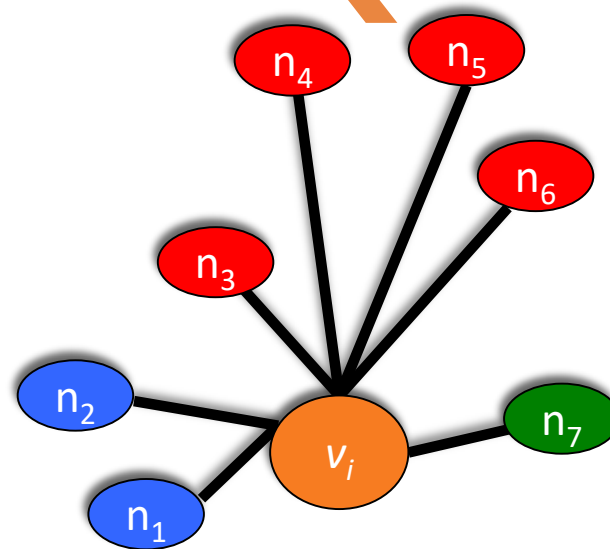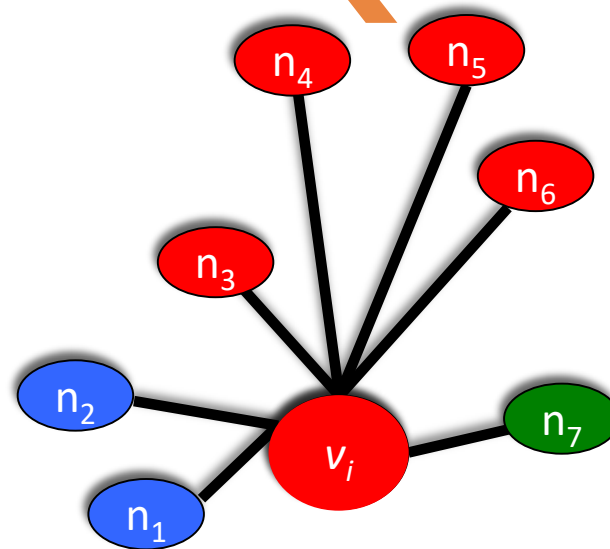| Colors	Units
Red	■■■■■■■
Green	■
Blue	■■

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.
2. Decide membership by selecting the dominant color.
3. Send "**all**" of received color amount back to the neighbors.

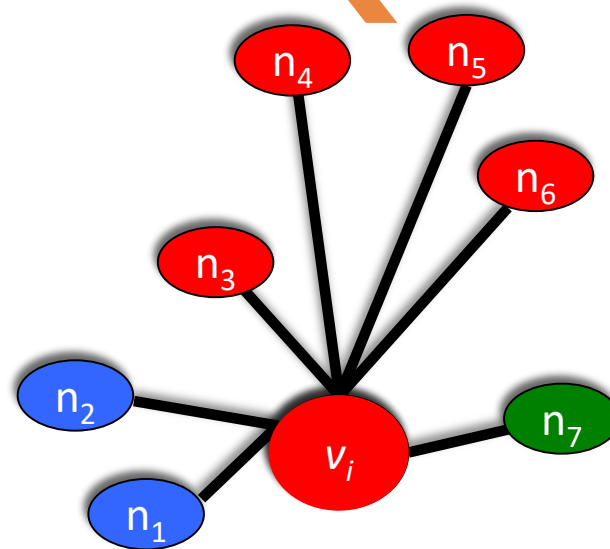| Colors	Units
Red	■■■■■■■
Green	■
Blue	■■

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.
2. Decide membership by selecting the dominant color.
3. Send "**all**" of received color amount back to the neighbors.

Colors   **Units**

Red

Green

Blue

Diffusion @ v_i

Regular diffusion steps:

1. Receive colors form neighbors.
2. Decide membership by selecting the dominant color.
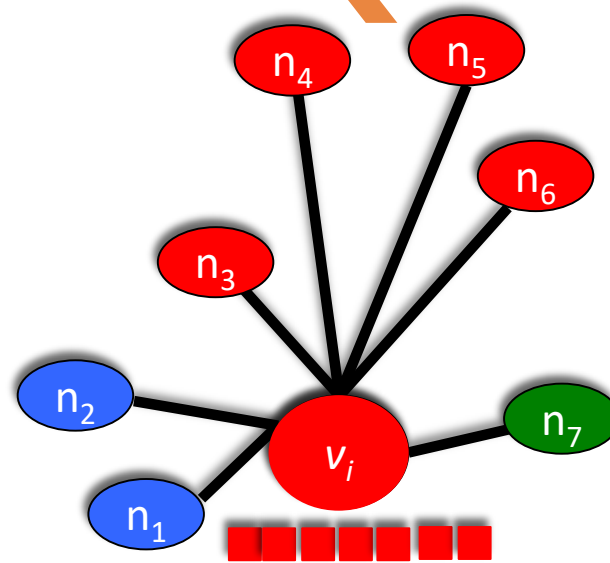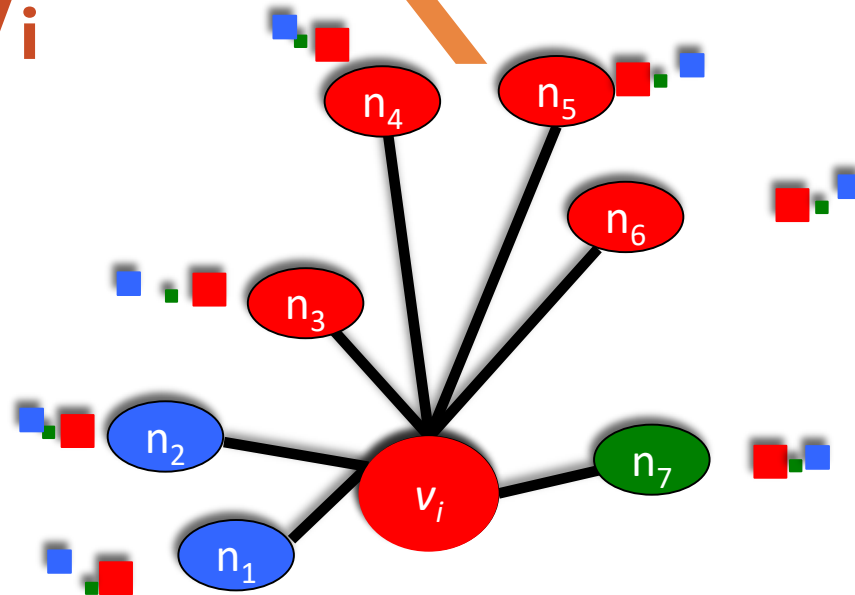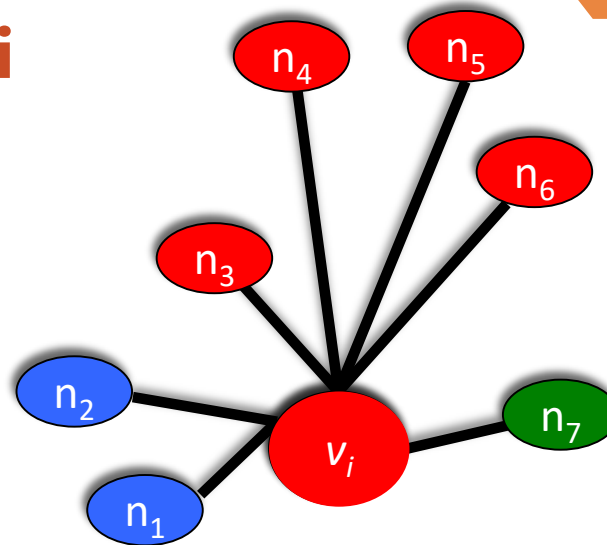3. Send "**all**" of received color amount back to the neighbors.

Colors	Units
Red	
Green	
Blue	

Stateful Diffusion (Stad) @ v_i

Stad steps:

1. Receive colors form neighbors.

2. Decide membership by selecting the dominant color.

3. Slow down diffusion speed: send **"some"** of received color amount back to the neighbors.

Colors	Units
Red	🟥🟥🟥🟥🟥🟥🟥
Green	🟩
Blue	🟦🟦

Stateful Diffusion (Stad) @ v_i

Stad steps:

1. Receive colors form neighbors.

2. Decide membership by selecting the dominant color.

3. Slow down diffusion speed: send **"some"** of received color amount back to the neighbors.

Colors   **Units**

Red

Green

Blue

Stateful Diffusion (Stad) @ v_i

Stad steps:

1. Receive colors form neighbors.

2. Decide membership by selecting the dominant color.

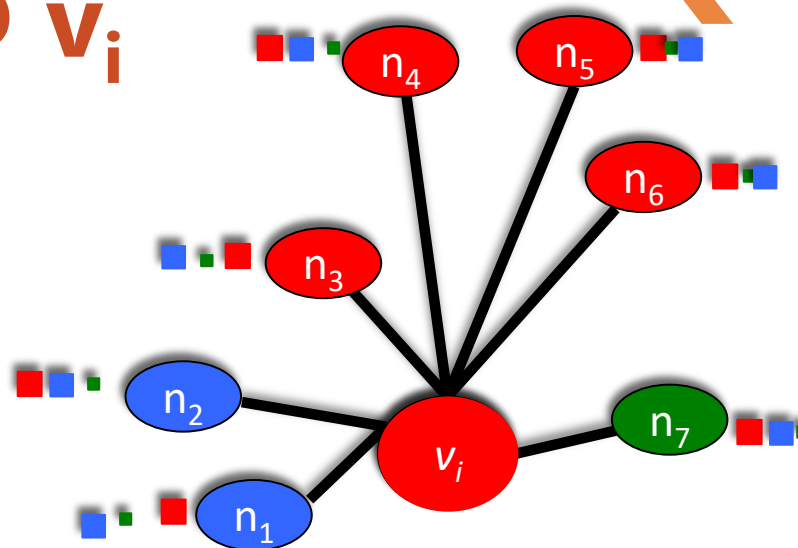3. Slow down diffusion speed: send **"some"** of received color amount back to the neighbors.
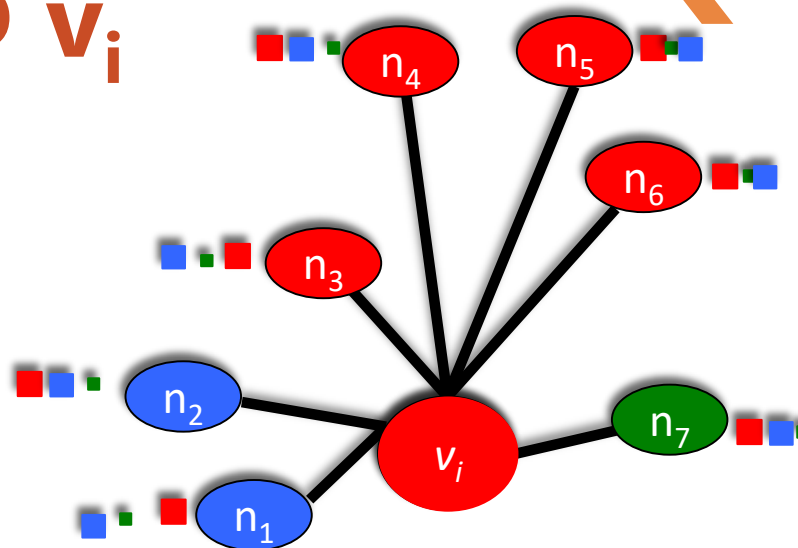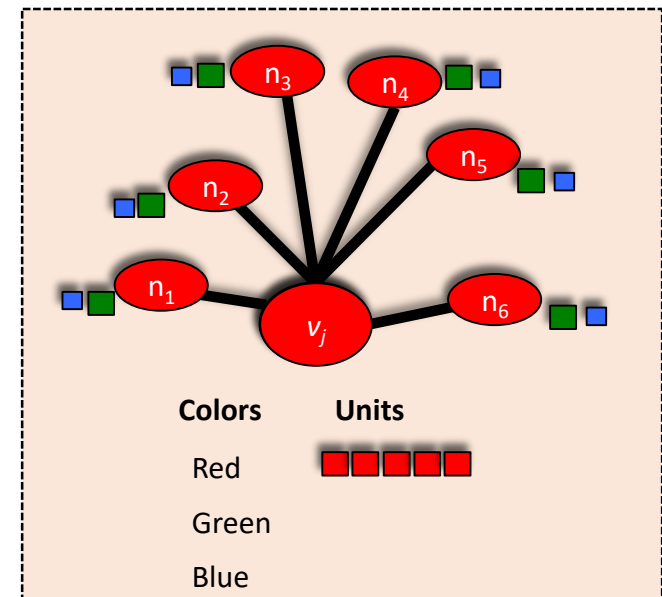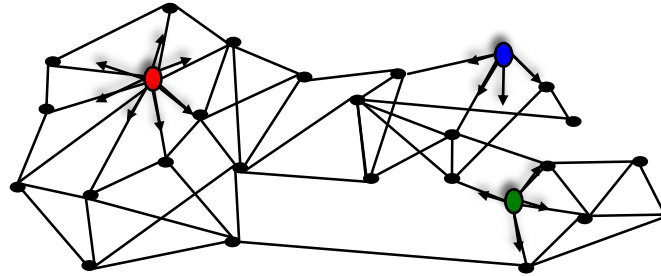
Stad: Fixed vs Adaptive Diffusion

(a) Initialization by selecting a set of random seeds

Iterations:

(1)

(2)

(3)

(b) Diffusion

(c) Stad

Stad: Fixed vs Adaptive Diffusion

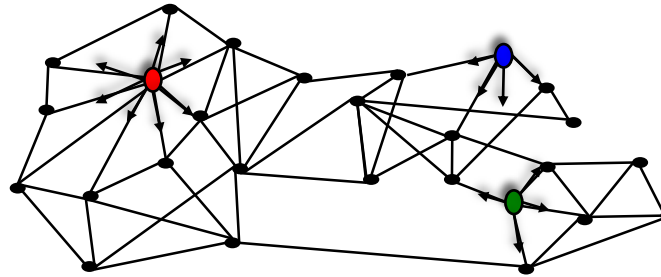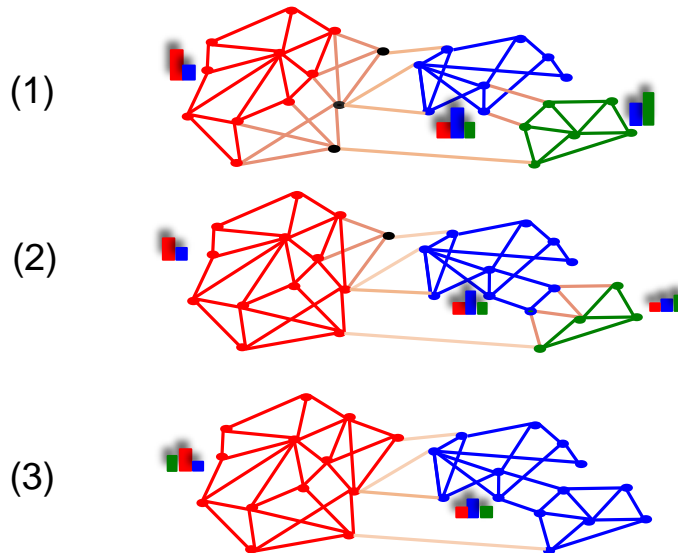(a) Initialization by selecting a set of random seeds

Iterations:

(1)

(2)

(3)

(b) Diffusion

(c) Stad

Stad: Fixed vs Adaptive Diffusion

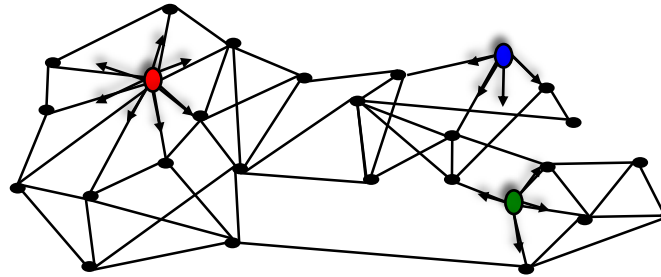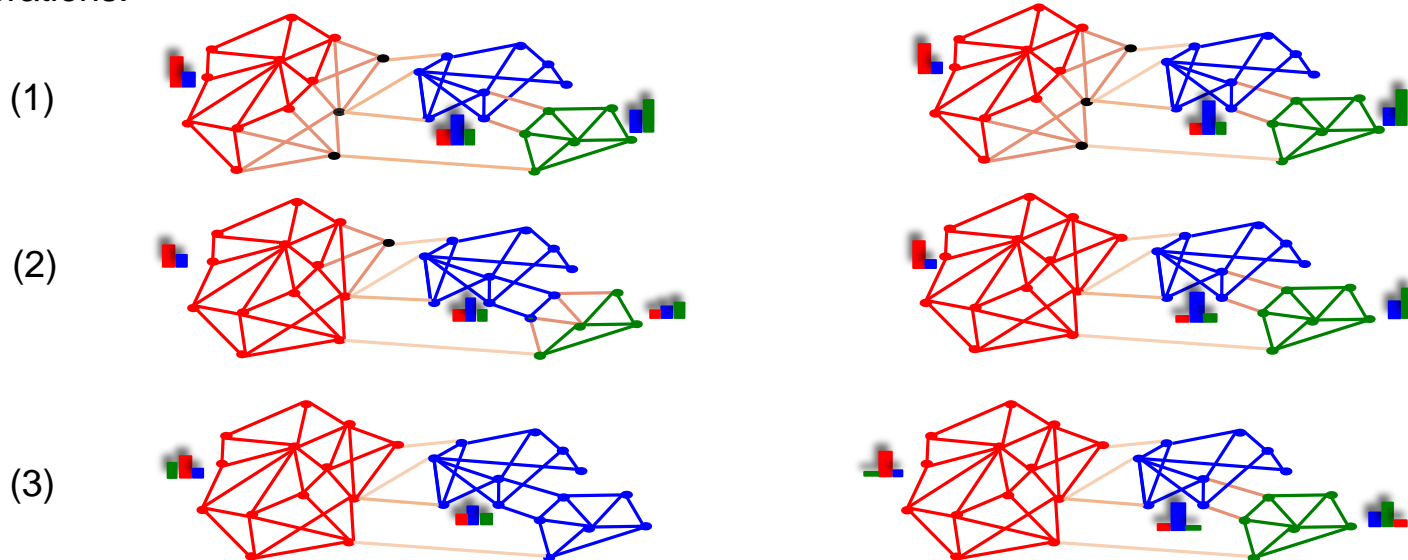**(a) Initialization by selecting a set of random seeds**
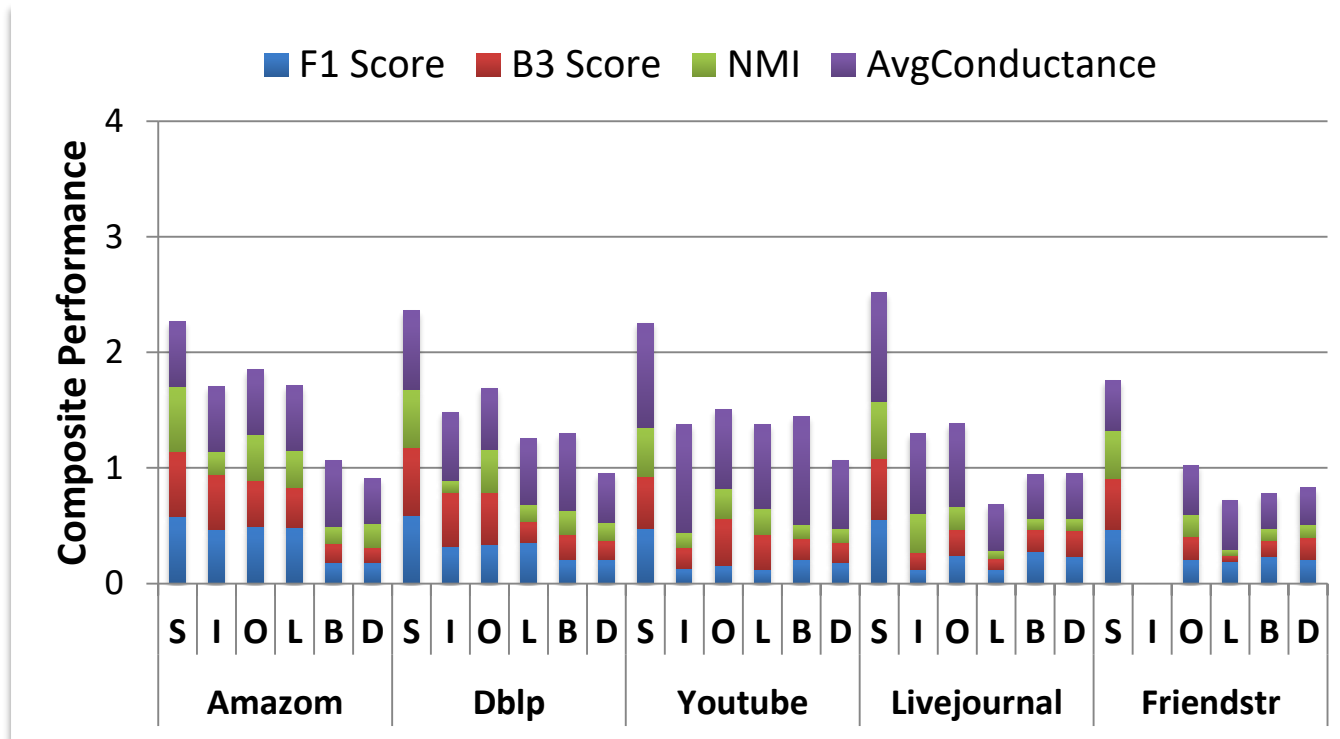
Iterations:

(1)

(2)

(3)

(b) Diffusion

(c) Stad

Experimental Results

Accuracy compared to baseline approaches

Performance comparison with the-sate-of-the-art community detection approaches:
Stad (S), Infomap (I), Oslom (O), Louvain (L), Bigclam (B), Diffusion (D).

Conclusion

**Decentralized Community Detection**

Stateful Diffusion

- Extracting communities with heterogenous size distribution.
- Performing optimization at node level while handling each community flow independently.
- Extracting disjoint as well as multiple community membership(s).

Future Work

- Adapting Stad to directed graphs that would require handling the challenges of edge directionality (dead-ends, cycles, etc.)

Experimental Results

Stad: Selecting seeds for color initialization
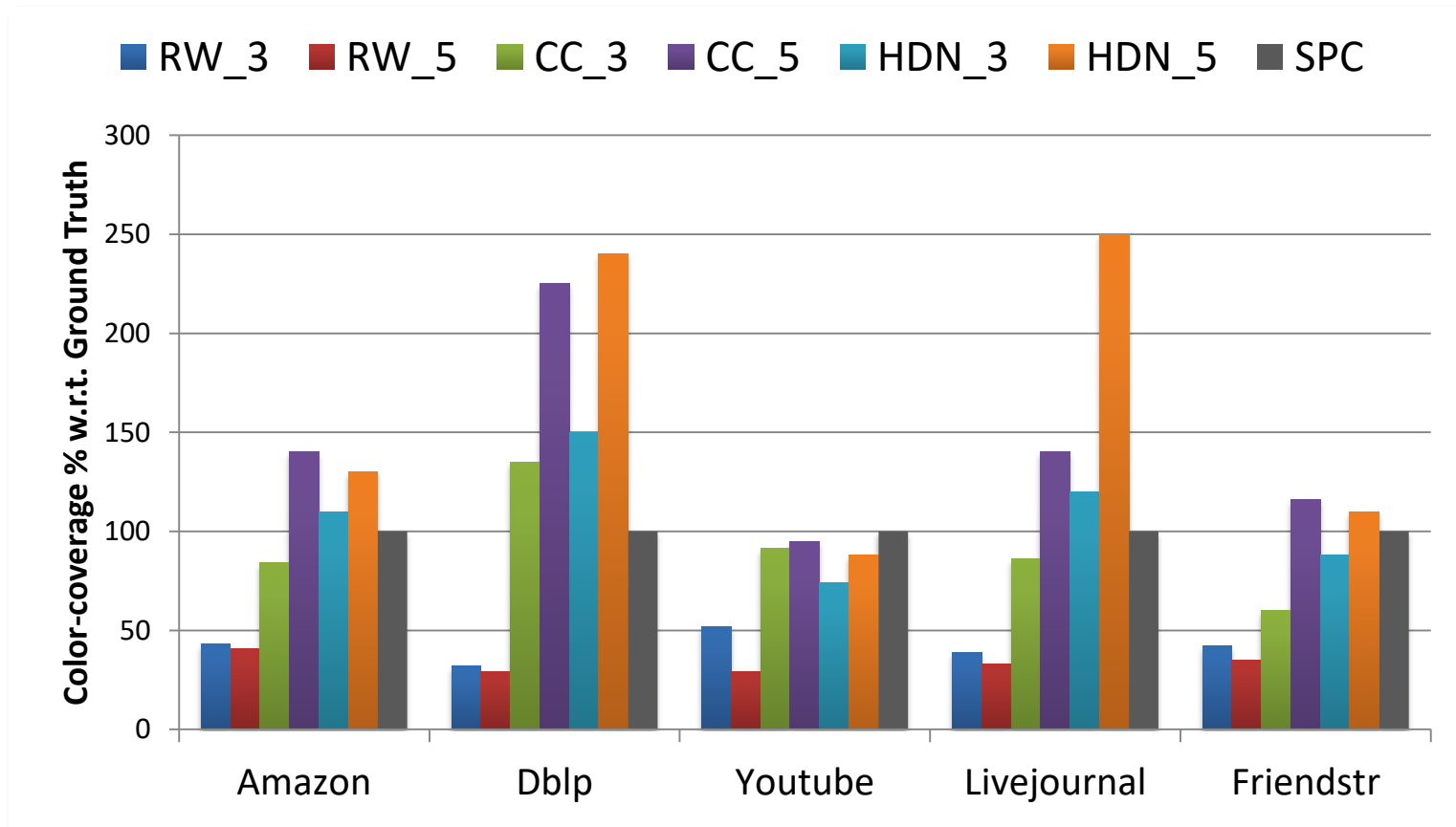
Thanks and questions are very welcomed ☺