

A First Course in Network Science

Chapter 6.1

Cunliang Geng
Amsterdam, 07/06/2021



▼ 6 Communities **clusters modules**

6.1 Basic Definitions

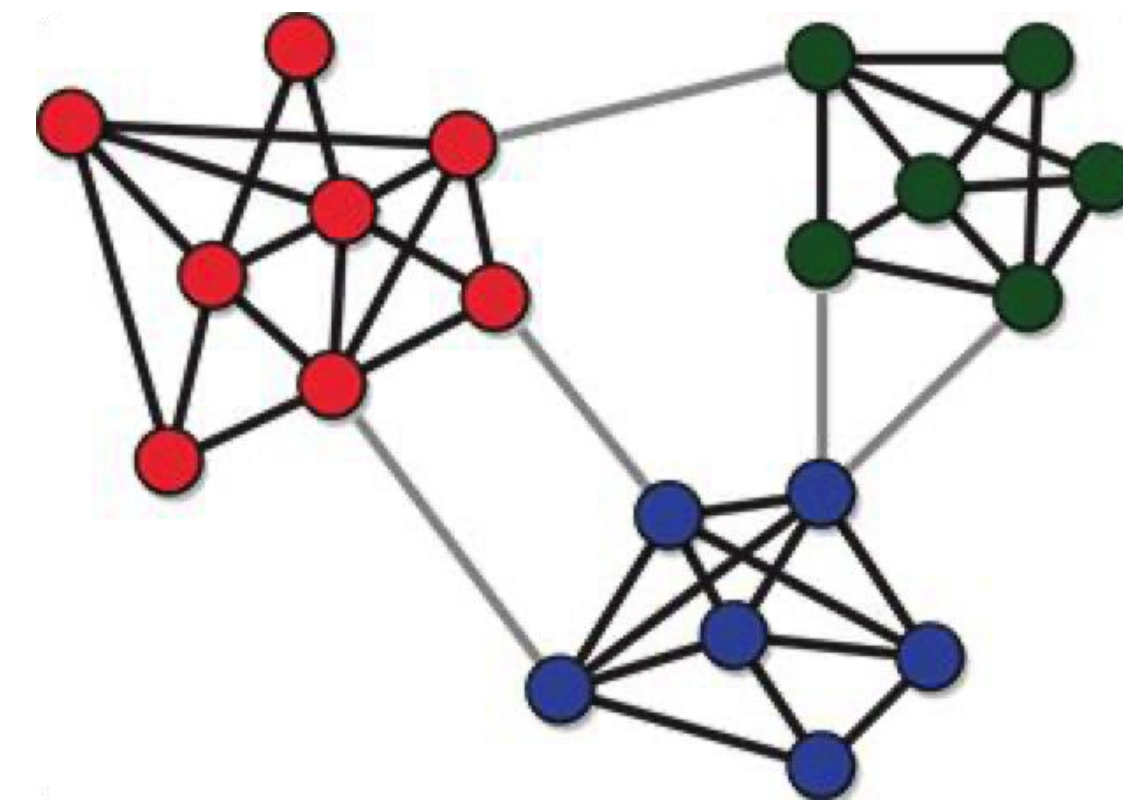
6.2 Related Problems

6.3 Community Detection

6.4 Method Evaluation

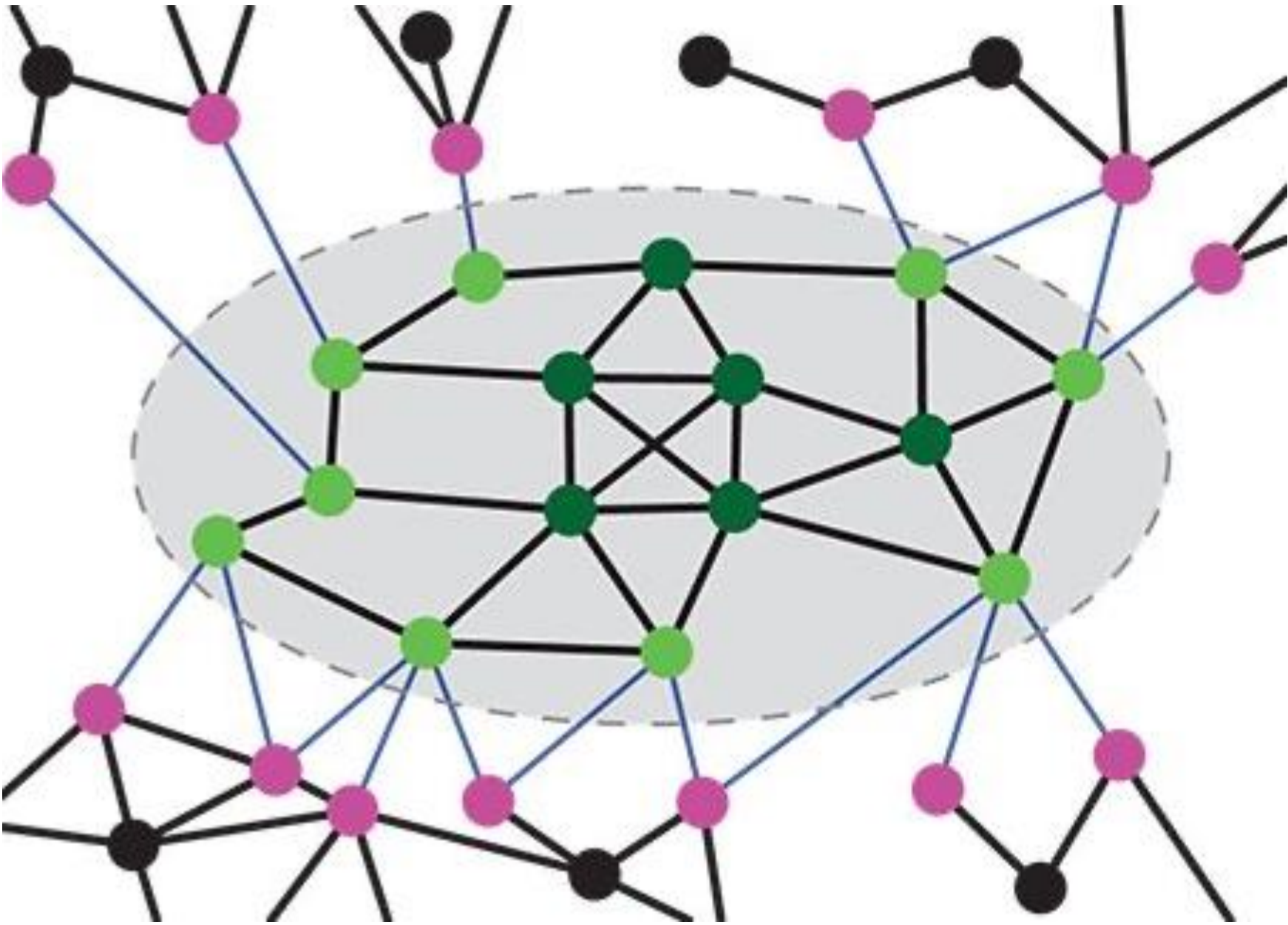
6.5 Summary

- How to describe a community
- How to define a community
- The partitions of community



How to describe a community

The community C		The oval part in the figure
The number of nodes	N_c	The green nodes in the oval
The number of internal links	L_c	The black lines in the oval
Internal degree of node i	k_i^{int}	The number of neighbors inside the oval, i.e. #black lines
External degree of node i	k_i^{ext}	The number of neighbors outside the oval, i.e. #blue lines
Degree of node i	$k_i = k_i^{int} + k_i^{ext}$	$k_i^{int} > 0 \ \& \ k_i^{ext} = 0$, dark green nodes
		$k_i^{int} > 0 \ \& \ k_i^{ext} > 0$, light green and magenta nodes
		$k_i^{int} = 0 \ \& \ k_i^{ext} > 0$, black green nodes
The community degree	$k_c = \sum_{i \in C} k_i$	The sum of degrees of the nodes in the oval
Internal link density	$\delta_c^{int} = \frac{L_c}{\binom{N_c}{2}} = \frac{2L_c}{N_c(N_c - 1)}$	The ratio between the #(internal links) and the max number of links that could exist between any two community nodes



Undirected and unweighted network

How to define a community

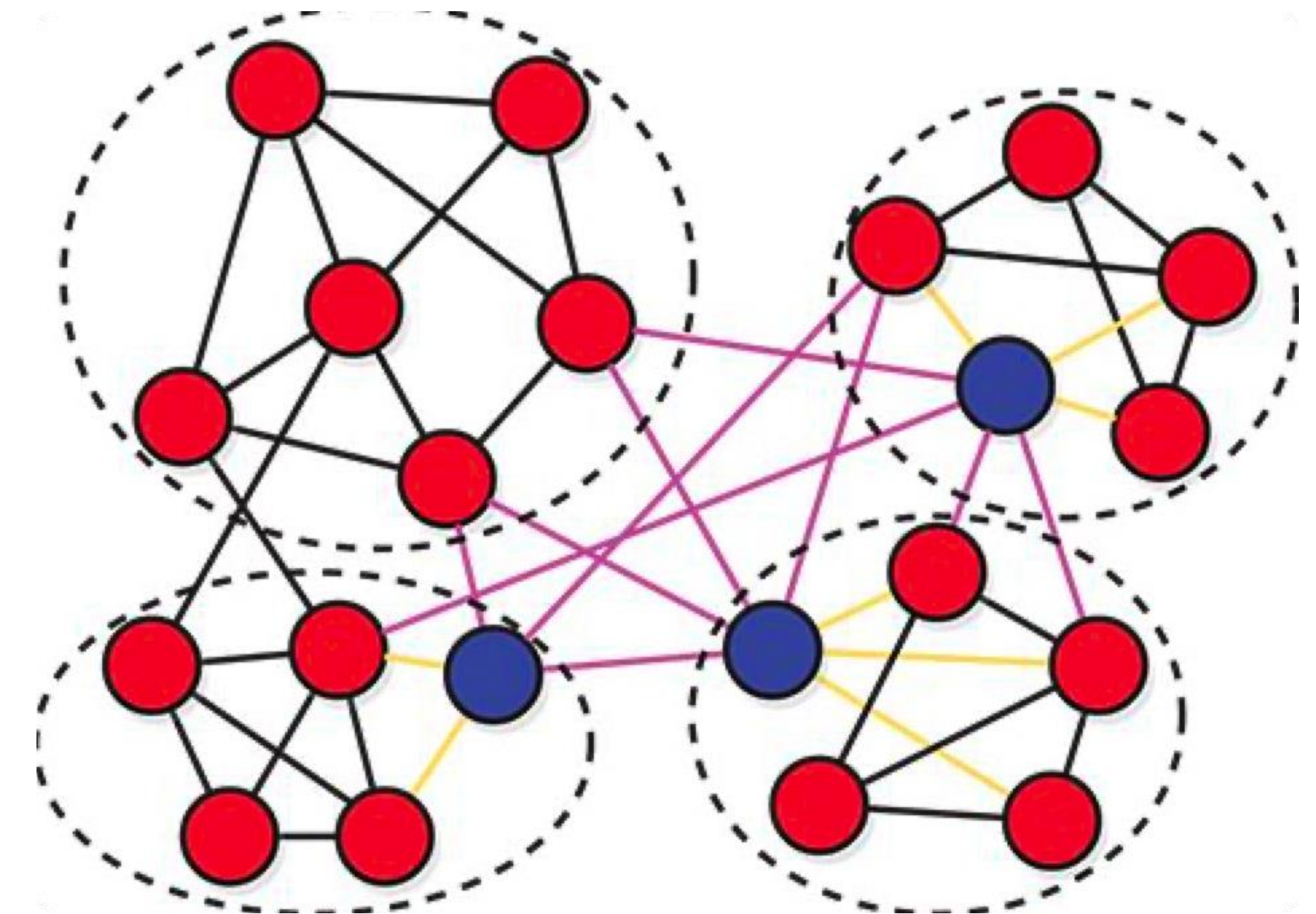
In general, traditional definitions rely on how to count internal and external links.

Treating the rest network of a community as a whole:

1. *Strong community*: a subnetwork such that the internal degree of each node exceeds its external degree.
2. *Weak community*: a subnetwork such that the sum of the internal degrees of all nodes exceeds the sum of their external degrees.

Treating the rest network as different subnetworks/communities:

1. *Strong community*: a subnetwork such that each node has more neighbours in it than in any other subnetworks.
2. *Weak community*: a subnetwork such that the sum of the internal degrees of all nodes exceeds the sum of their external degrees in each of the other subnetworks.



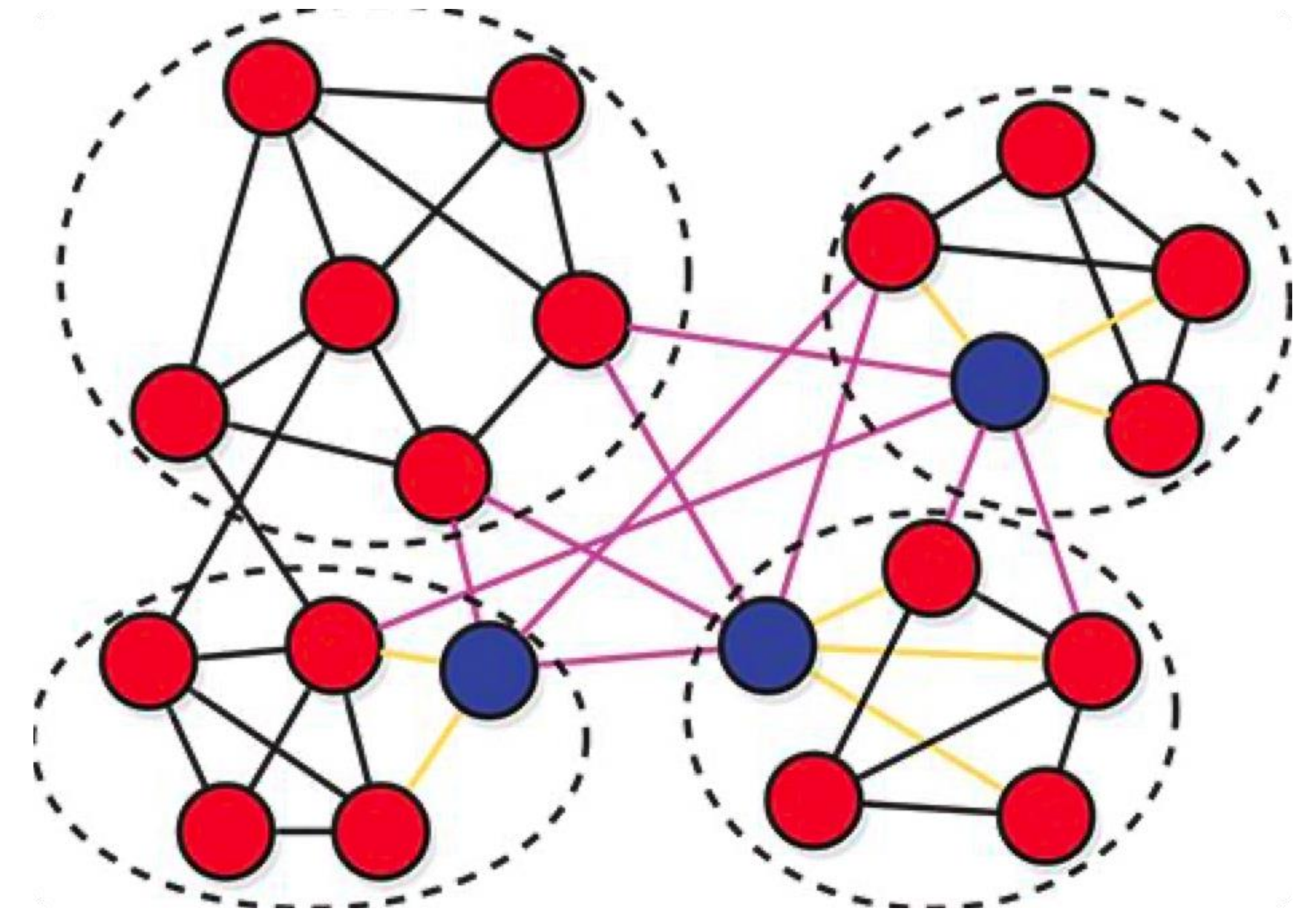
How to define a community

The problem of traditional definitions:

The number of links usually increases with the community size, so the comparisons between internal and external degrees are biased by their sizes.

Probability-based definition:

nodes within a community are more likely to be connected than nodes in different subnetworks.



Is a definition of community really necessary?

- Most network clustering methods do not require a precise definition of community.
- Defining criteria can be useful when checking the reliability of final results.

Partitions of community

A partition is a division, or grouping of a network into communities.

Bell number: the number of all possible partitions

Given a network with 15 nodes, its bell number is close to

- A. 100
- B. 1 000
- C. 1 000 000
- D. 1 000 000 000

It is 1 382 958 545

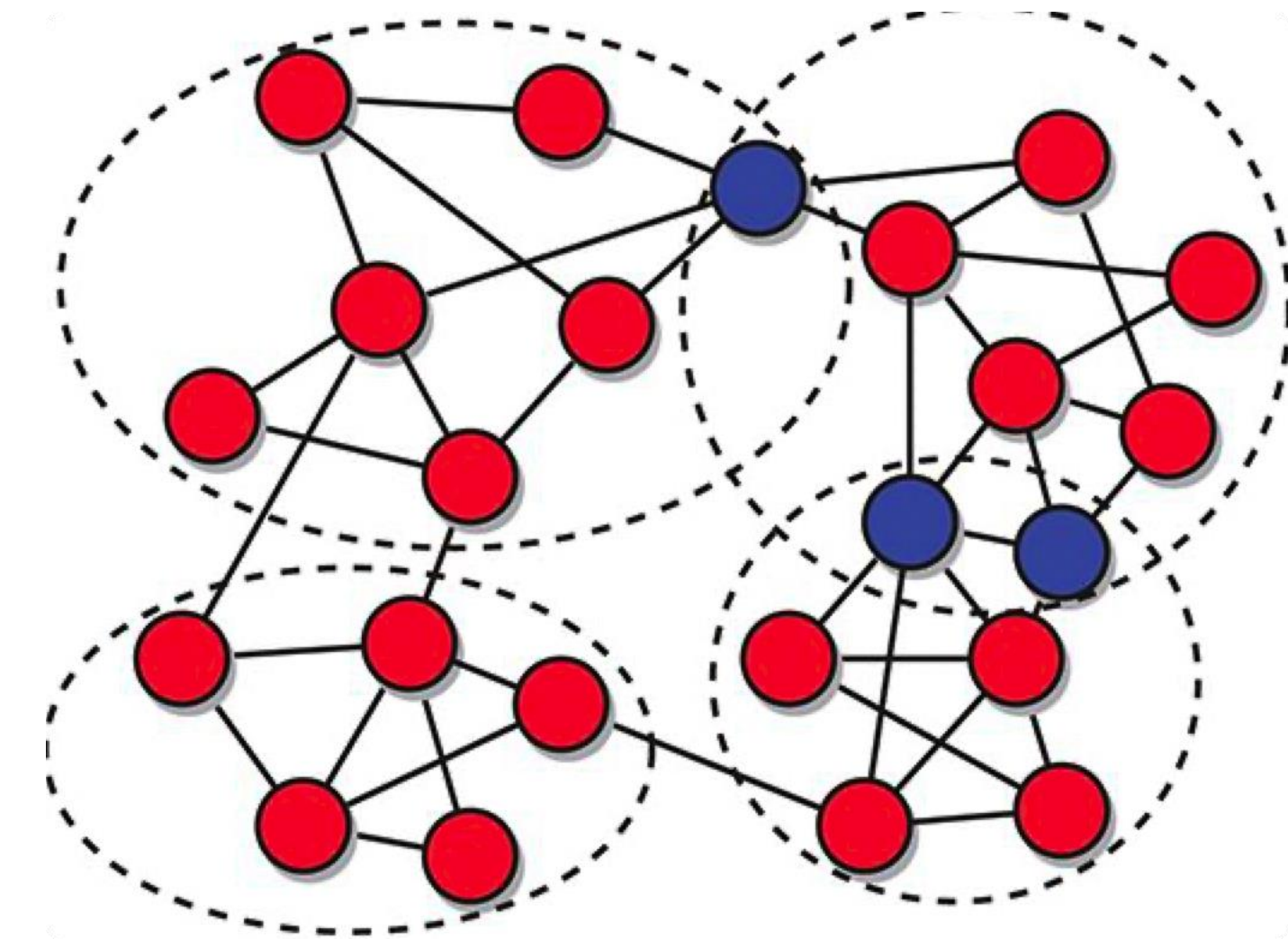
Do not try to choose the best partition by going through all possibilities

Partitions of community

The communities in many real networks overlap, i.e. they share some of their nodes

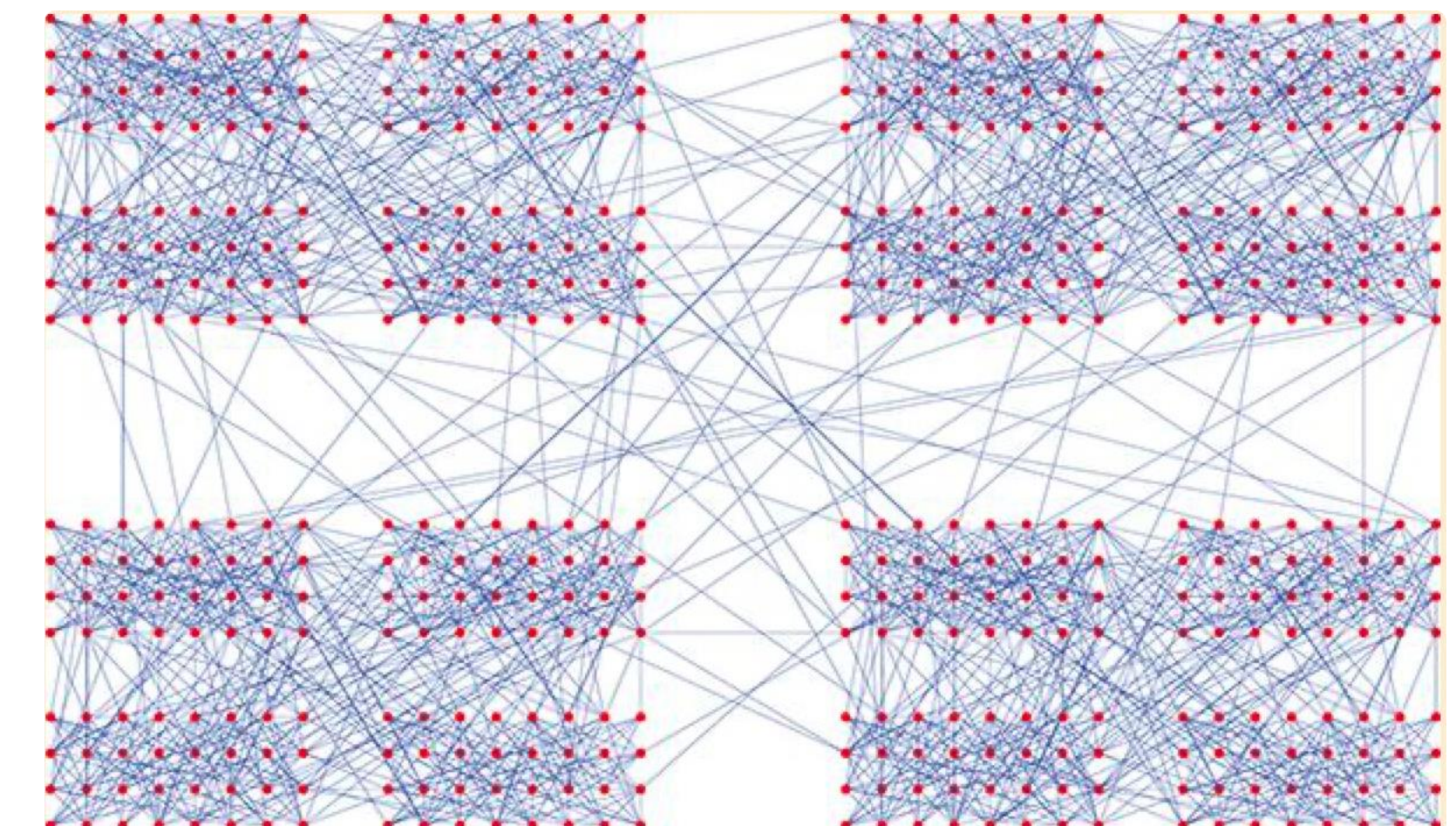
Cover: a division of network into overlapping communities

The number of possible covers \gg bell number



Partitions can be hierarchical when the network has multiple levels of organization at different scales

Partitions are often heterogeneous, some of the community properties may vary widely from one cluster to another



Thank you!
Q&A