Network Preliminaries

Héctor Corrada Bravo

University of Maryland, College Park, USA

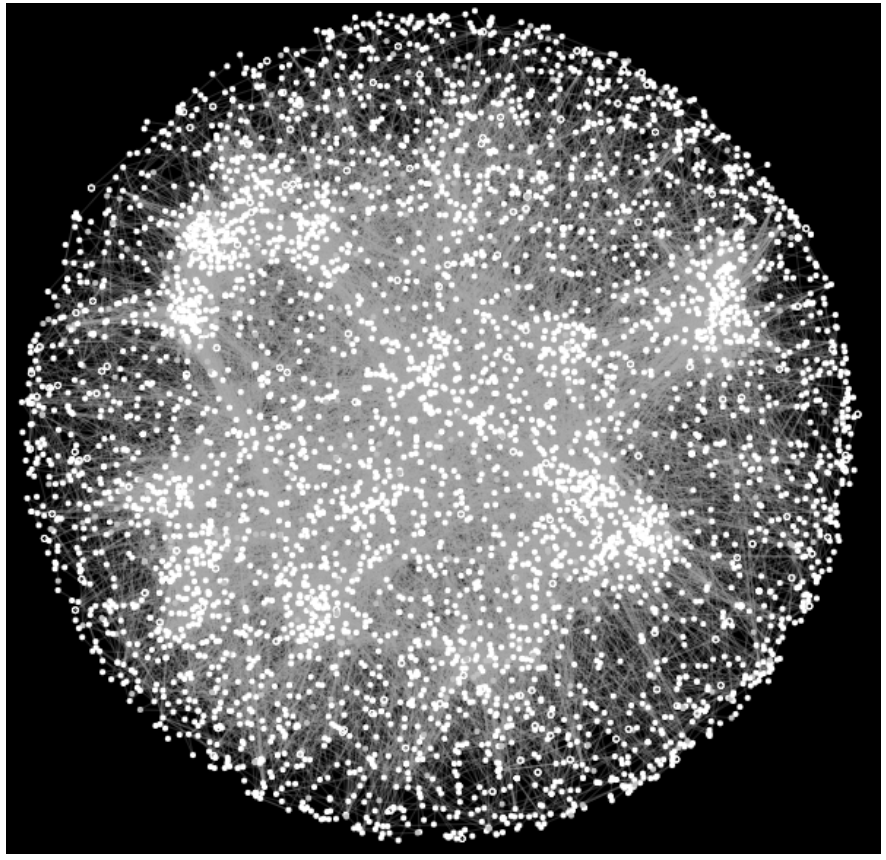
CMSC828O 2018-09-01

CBCB Center for Bioinformatics & Computational Biology
University of Maryland

Genetic Interaction Network

- Yeast high-throuput double-knockdown assay
- ~5000 genes
- ~800k interactions

http://www.geneticinteractions.org/



Costanzo et al. (2016) Science. DOI: 10.1126/science.aaf1420

Genetic Interaction Network

- Yeast high-throuput double-knockdown assay
- ~5000 genes
- ~800k interactions

http://www.geneticinteractions.org/

Costanzo et al. (2016) Science. DOI: 10.1126/science.aaf1420

Genetic Interaction Network

- Number of vertices: 2803
- Number of edges: 67,268

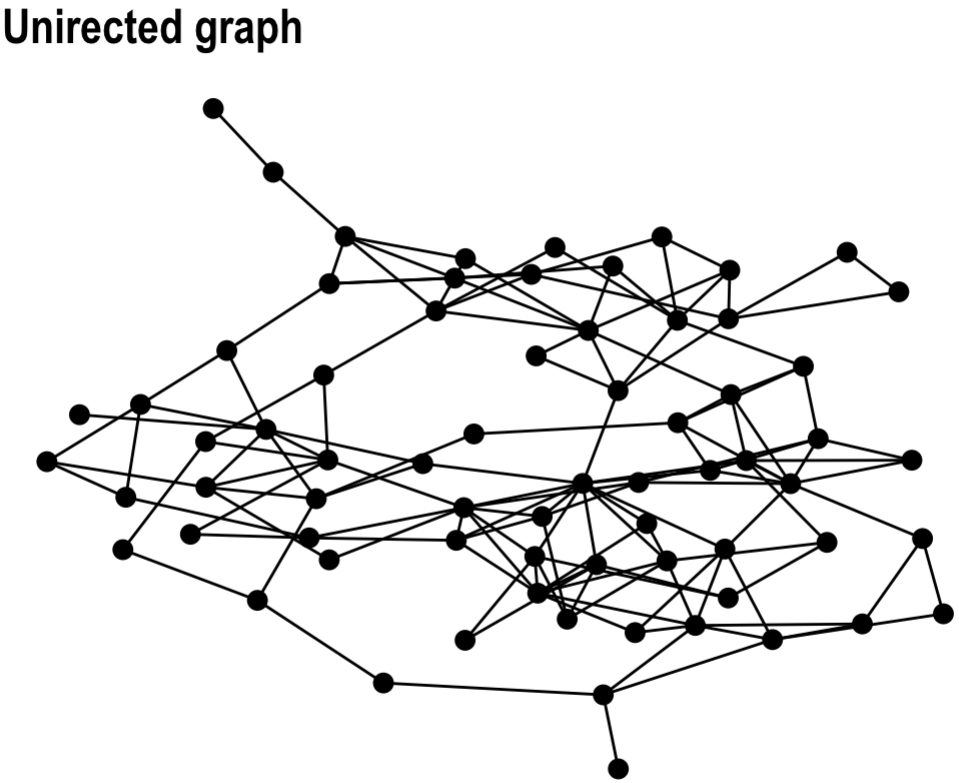
Preliminaries

Network: abstraction of *entities* and their interactions

Graph: mathematical representation

vertices: nodes

edges: links



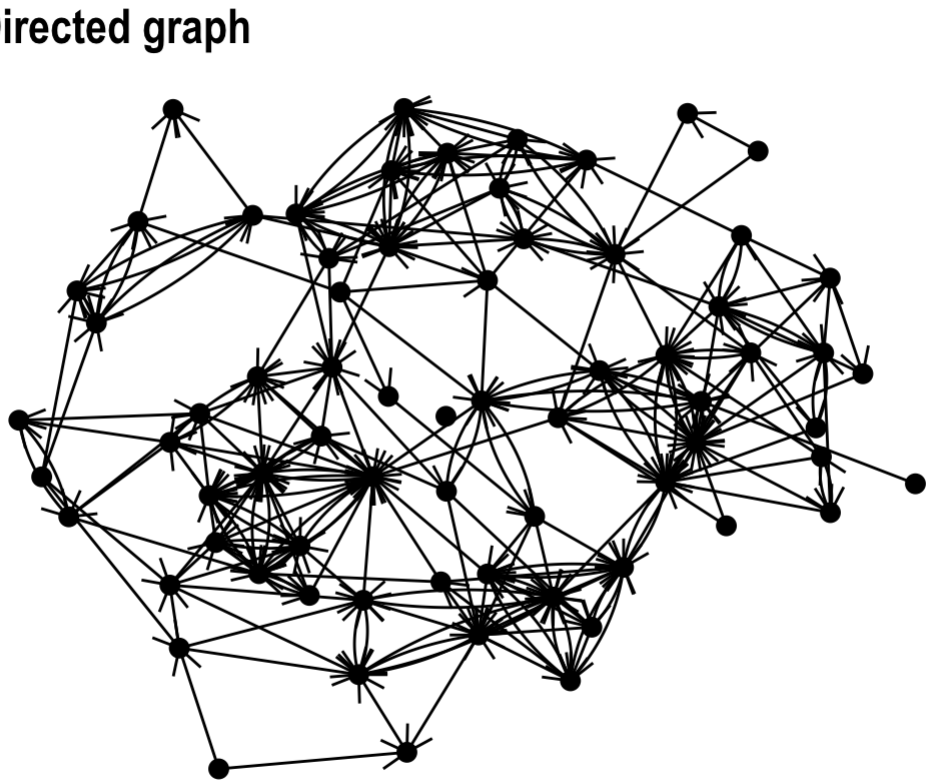
Preliminaries

Network: abstraction of *entities* and their interactions

Graph: mathematical representation

vertices: nodes

edges: links



Network statistics: notation

Number of vertices: n

In our example: *number of genes*

Network statistics: notation

Number of vertices: n

In our example: *number of genes*

Number of edges: m

In our example: *number of genetic interactions*

Network statistics: notation

Number of vertices: n

In our example: *number of genes*

Number of edges: m

In our example: *number of genetic interactions*

Degree of vertex i: k_i

Number of genetic interactions for gene i

Network statistics: notation

On the board:

- Calculate number of edges m using degrees k_i (for both directed and undirected networks)
- Calculate *average degree* c
- Calculate *density* ρ

Network statistics: notation

On the board:

- Calculate number of edges m using degrees k_i (for both directed and undirected networks)
- Calculate *average degree* c
- Calculate *density* ρ

In our example:

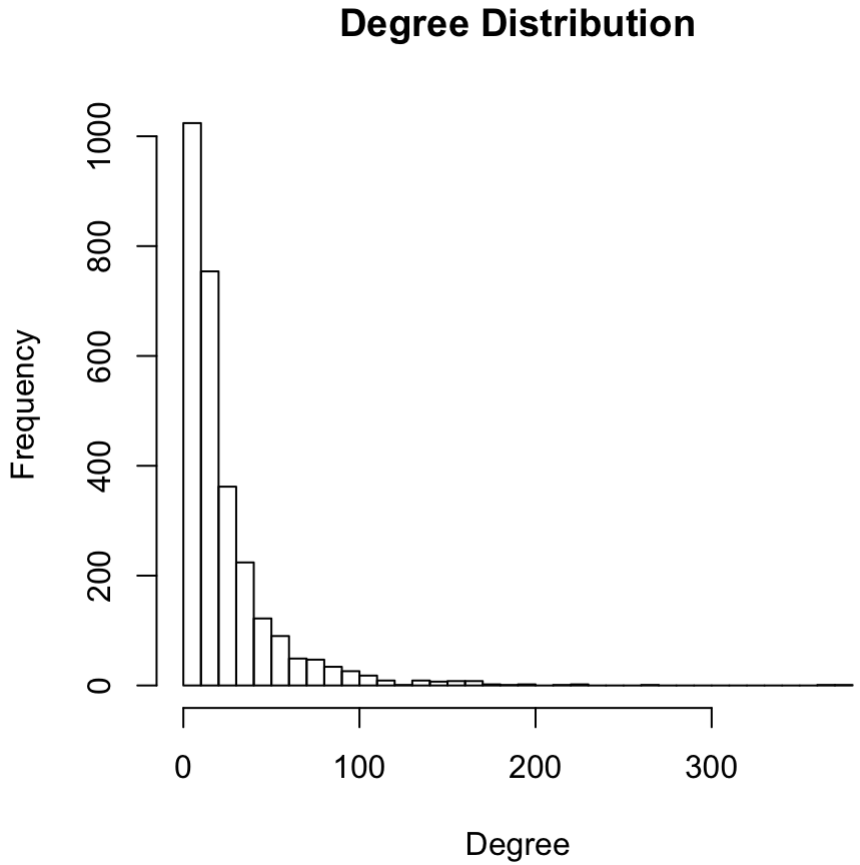
Average degree: 47.9971459
Density: 0.0171296

Degree distribution

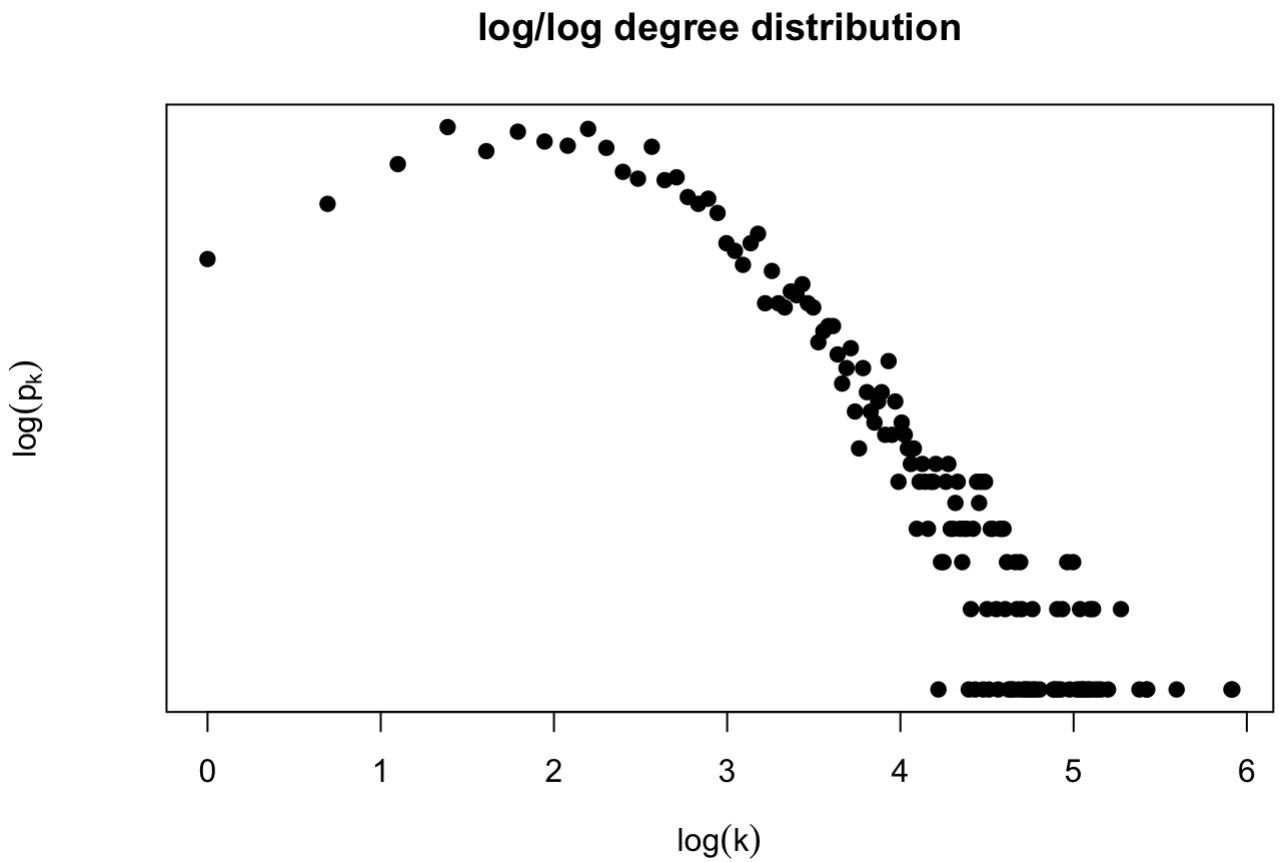
Fundamental analytical tool to characterize networks

p_k: probability randomly chosen vertex has degree k

On the board: how to calculate p_k and how to calculate average degree c using degree distribution.

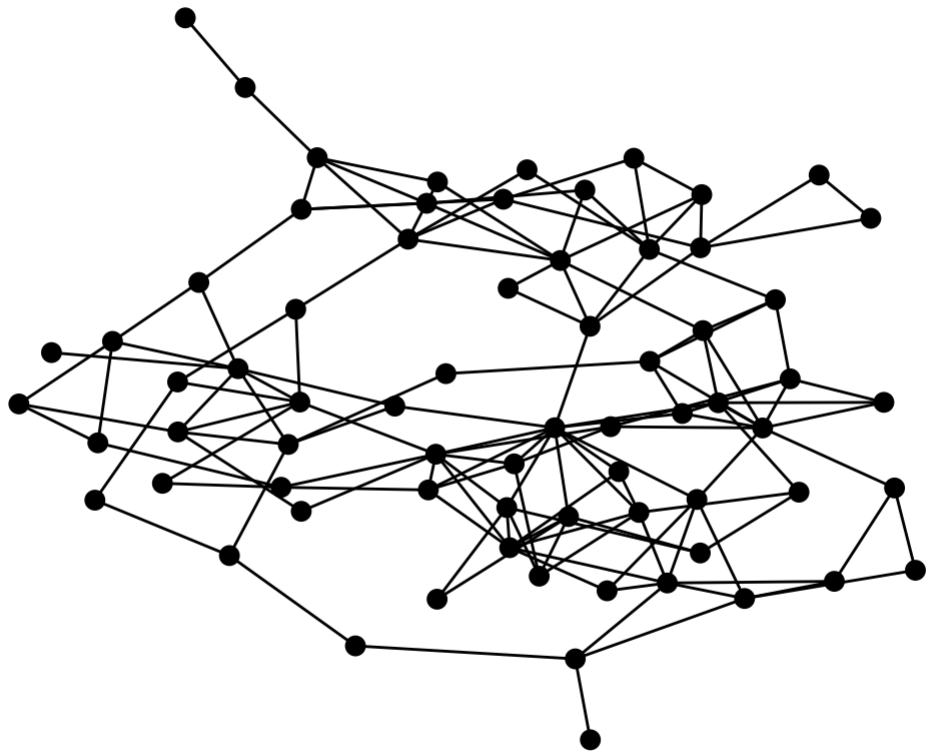


Degree Distribution



Paths and Distances

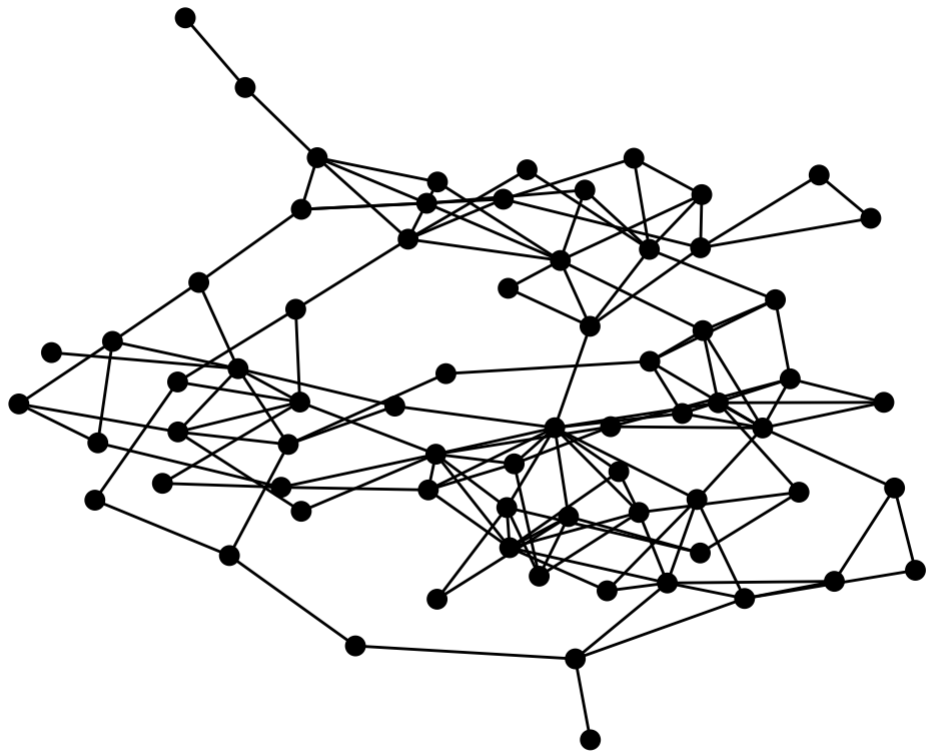
Distance d_{ij}: length of **shortest** path betwen vertices i and j.



Paths and Distances

Distance d_{ij}: length of **shortest** path betwen vertices i and j.

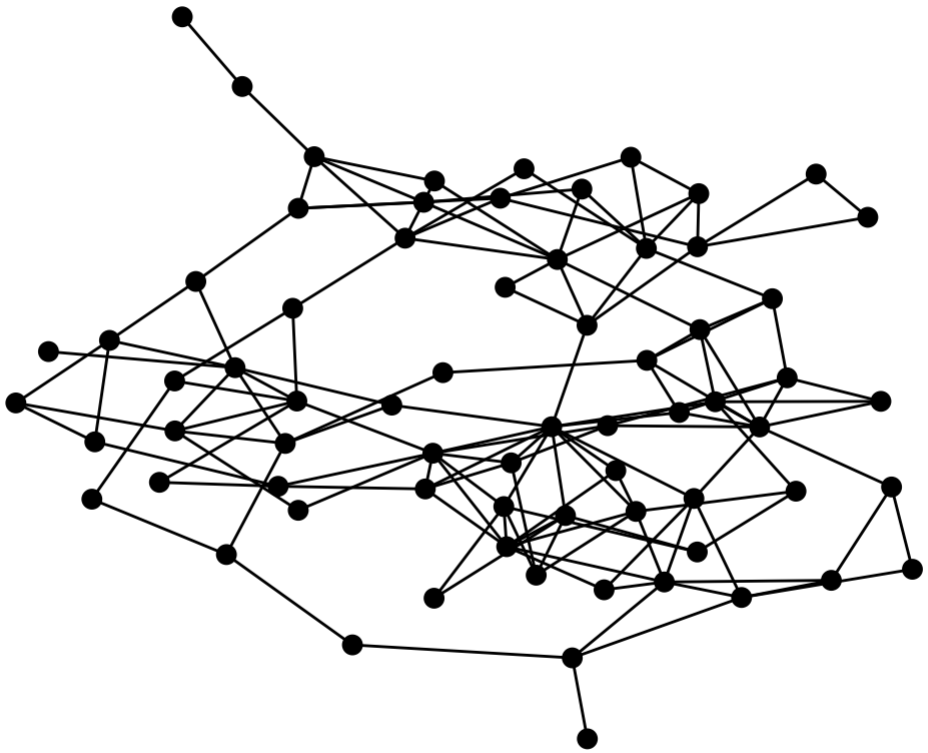
Diameter: longest shortest path $\max_{ij} d_{ij}$



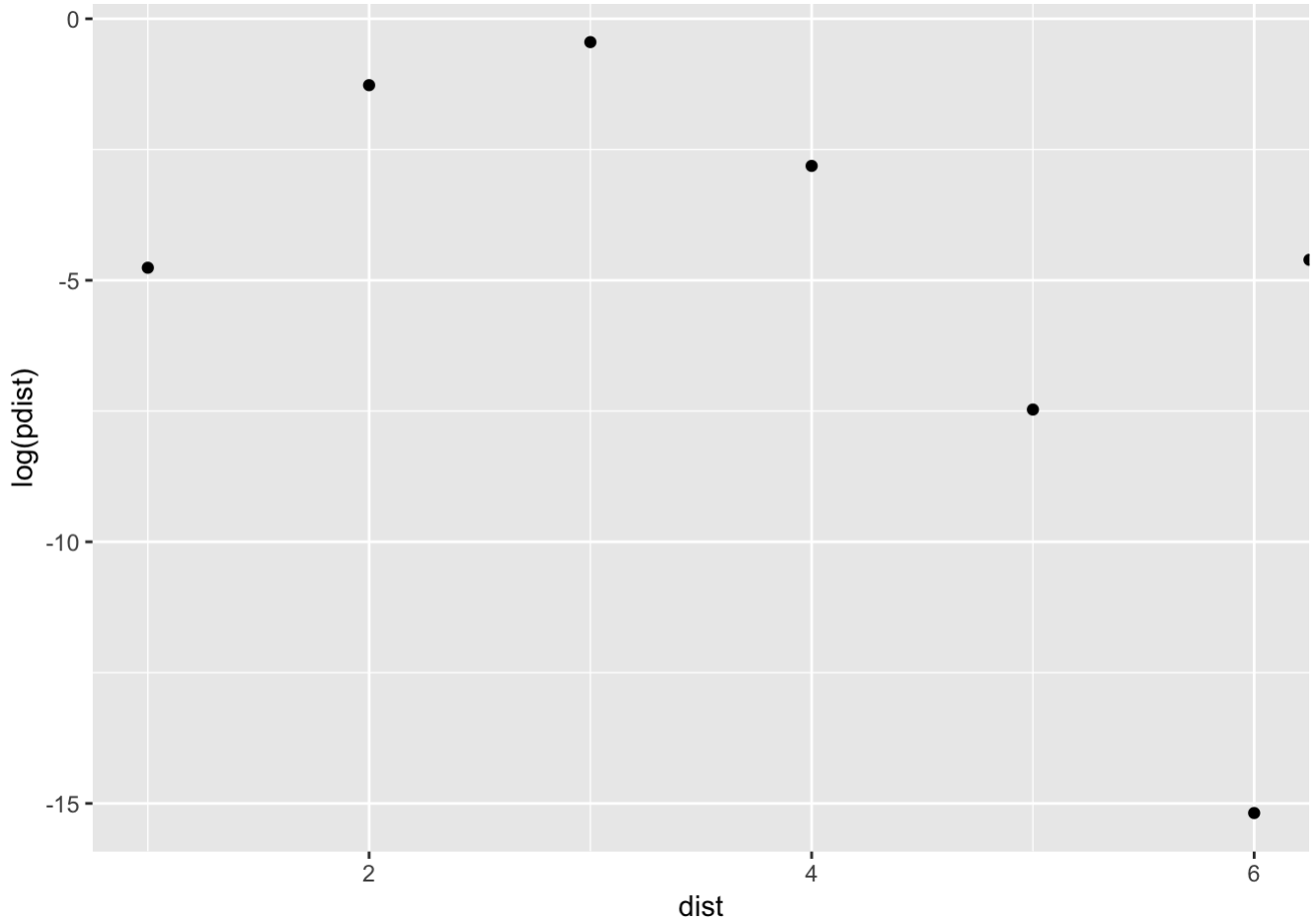
Paths and Distances

Distance d_{ij}: length of **shortest** path betwen vertices i and j.

On the board: average path length



Distance Distribution



Distances and paths

By convention: if there is no path between vertices i and j then $d_{ij} = \infty$

Distances and paths

By convention: if there is no path between vertices i and j then $d_{ij} = \infty$

Vertices i and j are *connected* if $d_{ij} < \infty$

Distances and paths

By convention: if there is no path between vertices i and j then $d_{ij} = \infty$

Vertices i and j are *connected* if $d_{ij} < \infty$

Graph is connected if $d_{ij} < \infty$ for all i, j

Distances and paths

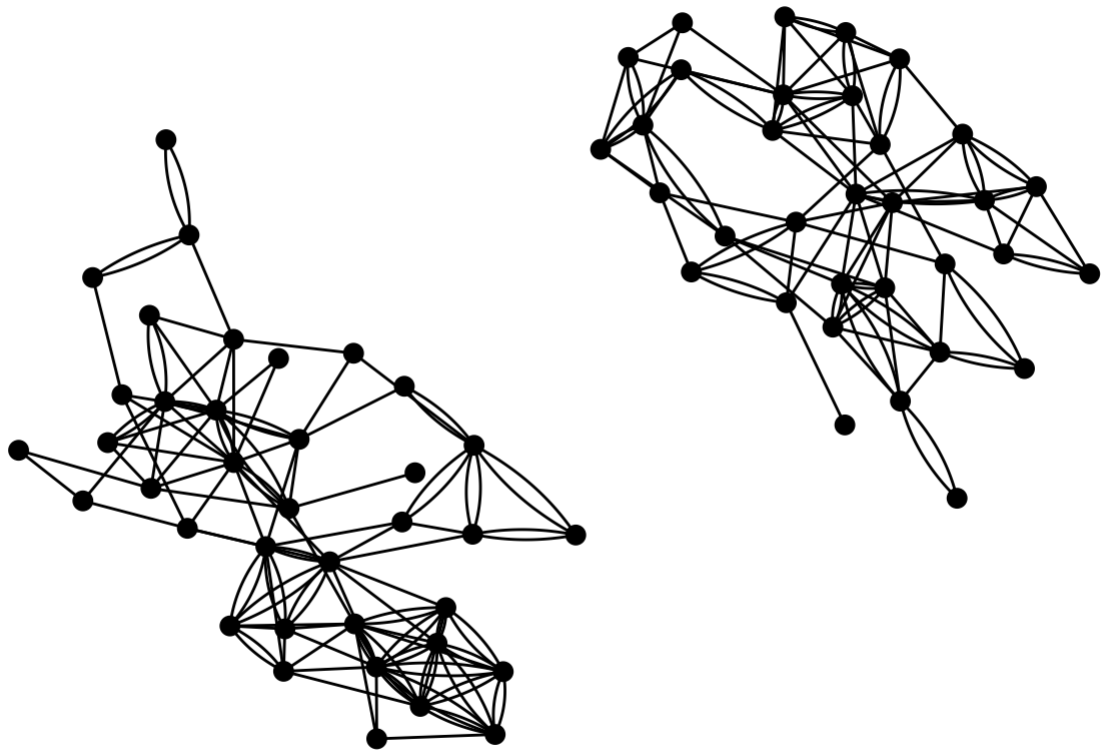
By convention: if there is no path between vertices i and j then $d_{ij} = \infty$

Vertices i and j are *connected* if $d_{ij} < \infty$

Graph is connected if $d_{ij} < \infty$ for all i, j

Components maximal subset of connected components

Components



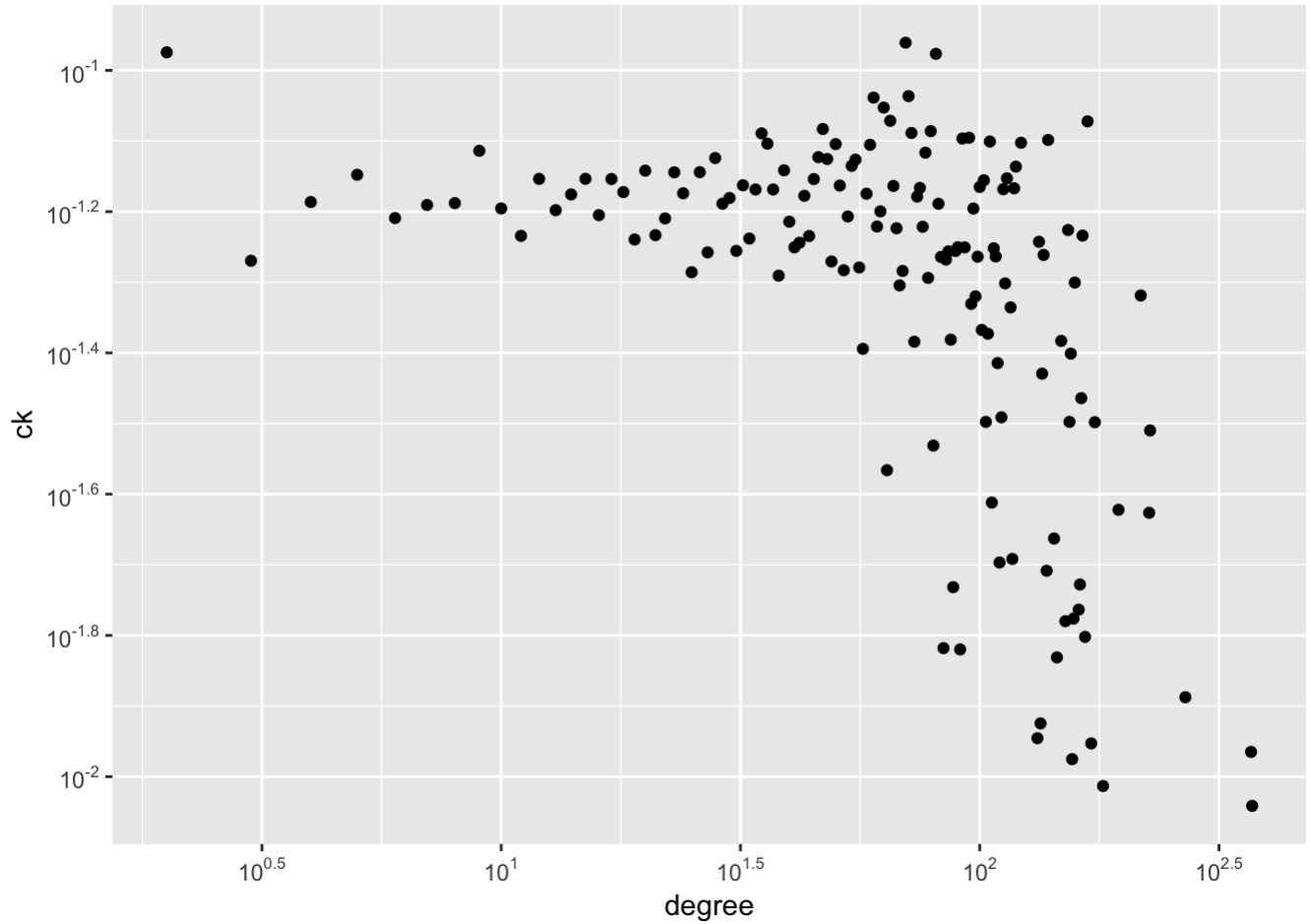
Clustering Coefficient

One last quantity of interest: how dense is the neighborhood around vertex i?

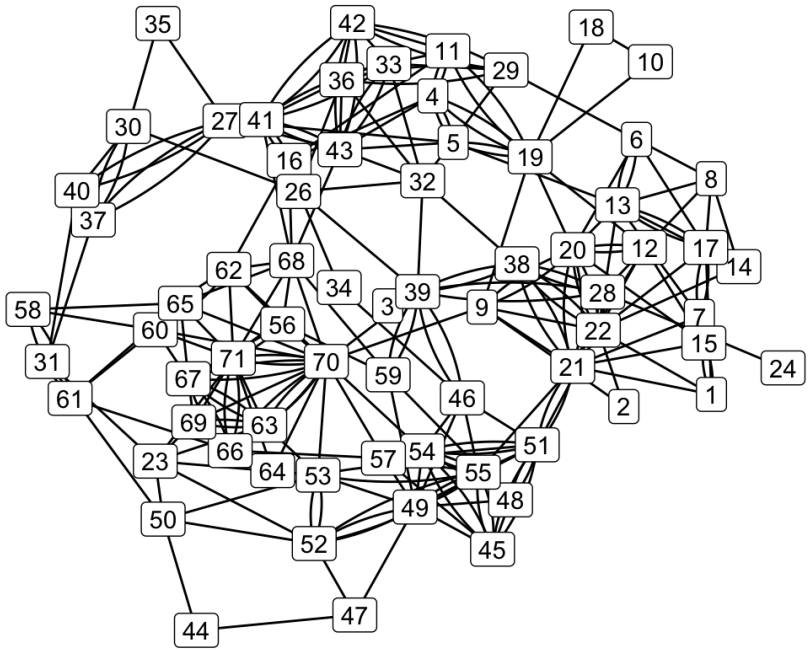
Do the genes that interact with me also interact with each other?

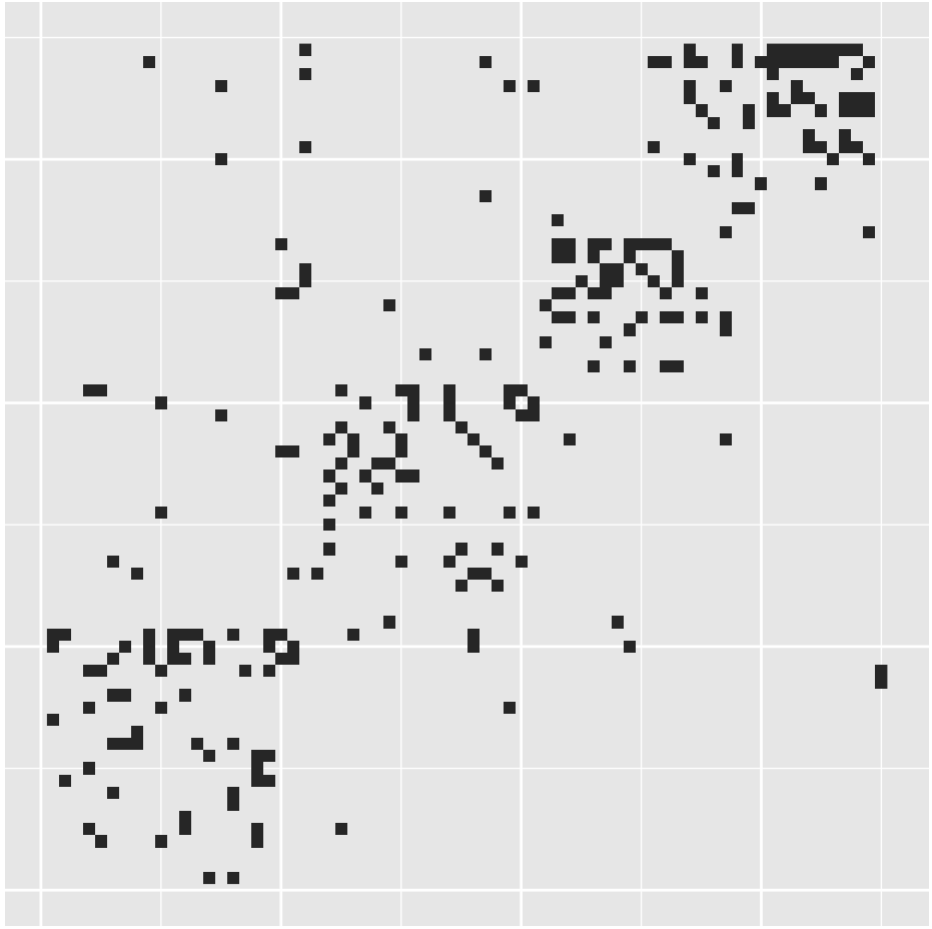
Definition on the board

Clustering coefficient



Adjacency Matrix





Adjacency Matrix

On the board:

- Definition
- Computing degree with adj. matrix
- Computing num. edges m with adj. matrix
- Computing paths with adj. matrix