

Part I

**CENTER FOR ANIMAL DISEASE MODELING AND SURVEILLANCE (CADMS),
SCHOOL OF VETERINARY MEDICINE, UC DAVIS**

**Jose Pablo Gomez, Jerome Baron, Jose Manuel Diaz Cao
Beatriz Martinez Lopez**

Center for Animal Disease Modeling and Surveillance (CADMS)
Department of Medicine & Epidemiology
School of Veterinary Medicine
University of California, Davis

- Emails: jpgo@ucdavis.edu
- jnbaron@ucdavis.edu

<https://jpablo91.github.io>

<https://cadms.vetmed.ucdavis.edu/>

Table of contents

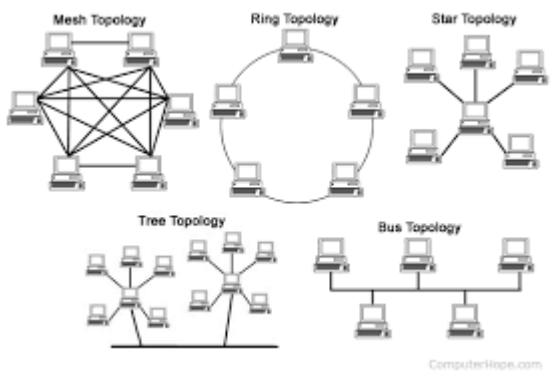
I. Introduction to networks

- A. What is a network?
- B. Elements of a network
- C. Data
 - a) Sources
 - b) Sampling

II. Static network analysis

- A. Network simplification
- B. Network properties
- C. Weak and Strong component

I.A - What is a network?



Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective risk-based targeted intervention**
 - **Surveillance:** Diagnostic testing, road checks



Pig sampling - Serbia

FAO African Swine Fever: Detection and diagnosis - 2017

Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective risk-based targeted intervention**
 - **Surveillance:** Diagnostic testing, road checks
 - **Prevention:** Education and information, vaccination campaigns



Farmer education – Burkina Faso

FAO African Swine Fever: Detection and diagnosis - 2017

Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective risk-based targeted intervention**
 - **Surveillance:** Diagnostic testing, road checks
 - **Prevention:** Education and information, vaccination campaigns
 - **Control:** Vaccination, treatment, culling campaigns, movement restrictions



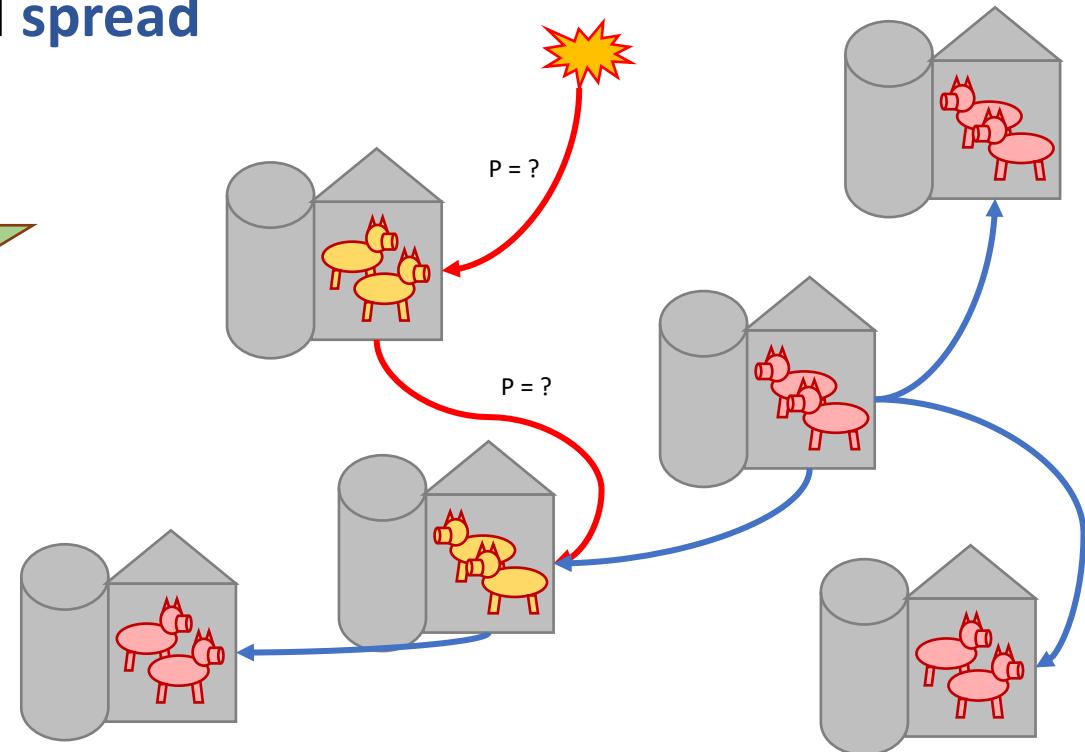
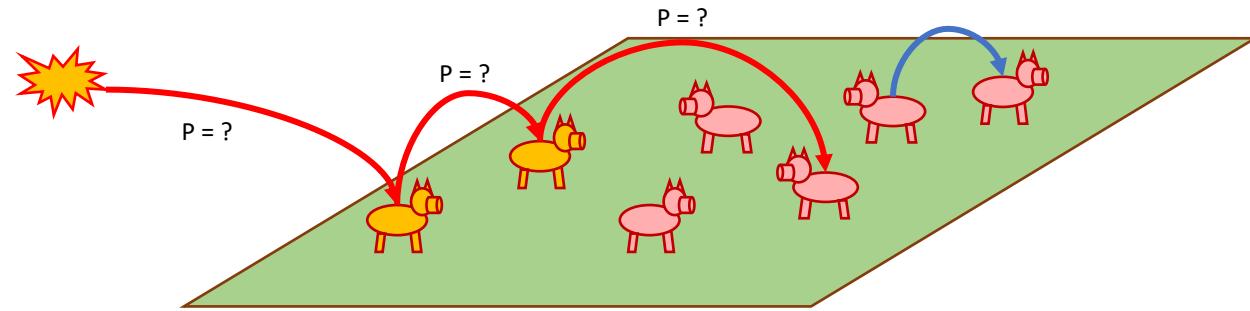
Road blocks and CO₂ culling chamber - Lithuania

FAO African Swine Fever: Detection and diagnosis - 2017

Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control

- Define strategic nodes for **cost-effective risk-based targeted intervention**
 - **Surveillance:** Diagnostic testing, road checks
 - **Prevention:** Education and information, vaccination campaigns
 - **Control:** Vaccination, treatment, culling campaigns, movement restrictions
- Modelling of disease **introduction and spread**

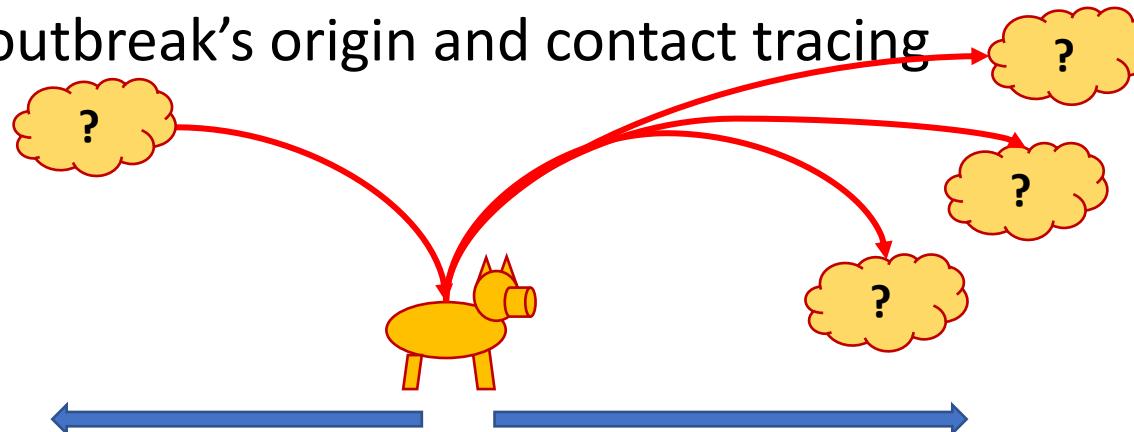


Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective risk-based targeted** intervention
 - **Surveillance:** Diagnostic testing, road checks
 - **Prevention:** Education and information, vaccination campaigns
 - **Control:** Vaccination, treatment, culling campaigns, movement restrictions
 - Modelling of disease **introduction and spread**
- **Risk Factor** Analysis

Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic nodes** for **cost-effective risk-based targeted intervention**
 - **Surveillance:** Diagnostic testing, road checks
 - **Prevention:** Education and information, vaccination campaigns
 - **Control:** Vaccination, treatment, culling campaigns, movement restrictions
 - Modelling of disease **introduction and spread**
- Risk Factor Analysis
- **Outbreak** investigation
 - Traceability of the outbreak's origin and contact tracing



Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective** surveillance
 - **Surveillance:** Diagnostic testing, road checks, monitoring
 - **Prevention:** Education and information, vaccination
 - **Control:** Vaccination, treatment, culling of infected individuals
 - Modelling of disease **introduction** and spread
- Risk Factor Analysis
- Outbreak investigation
 - Traceability of the outbreak's origin
- Other uses in the animal world
 - Behavioral studies in social animals: i.e. contact patterns in a herd of cattle, primate behavioral studies



Source: San Diego Zoo

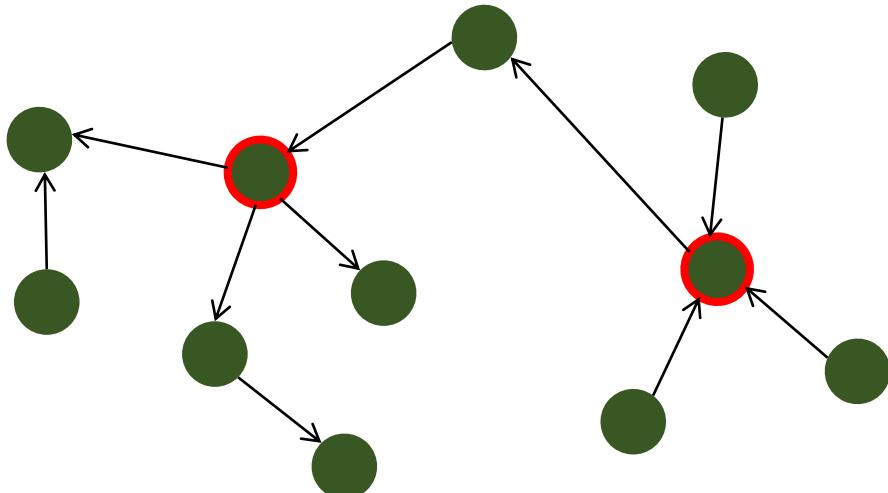
Applications in Preventative Veterinary Medicine

- Surveillance, Prevention and Control
 - Define **strategic** nodes for **cost-effective risk-based targeted** intervention
 - **Surveillance**: Diagnostic testing, road checks
 - **Prevention**: Education and information, vaccination campaigns
 - **Control**: Vaccination, treatment, culling campaigns, movement restrictions
 - Define **cost-effective risk-based targeted** approaches
 - Modelling of disease **introduction and spread**
- **Risk Factor** Analysis
- **Outbreak** investigation
 - Traceability of the outbreak's origin
- Other uses in the animal world
 - Behavioral studies in social animals: i.e. contact patterns in a herd of cattle

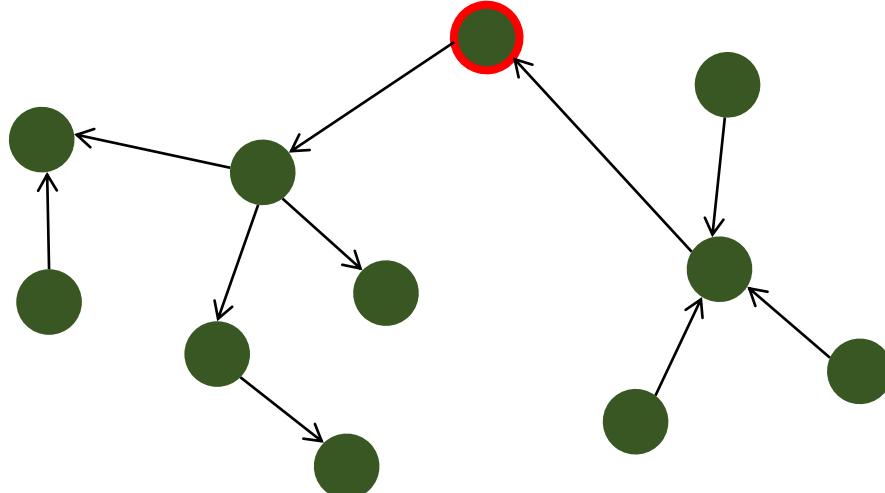
Why represent events in a network?

To describe contact dynamics

Identify individuals that are very active

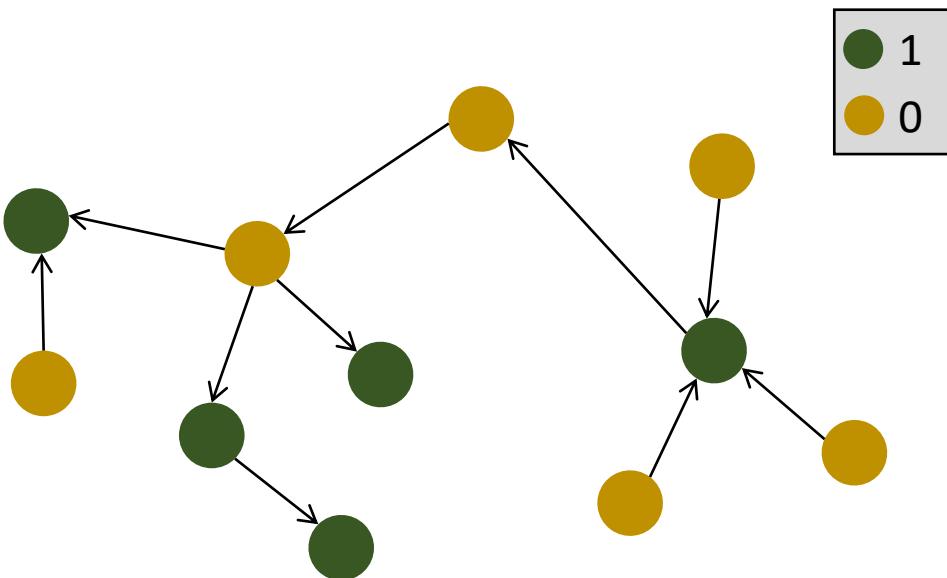


Identify individuals that are intermediate



Why represent events in a network?

Model contact dynamics:

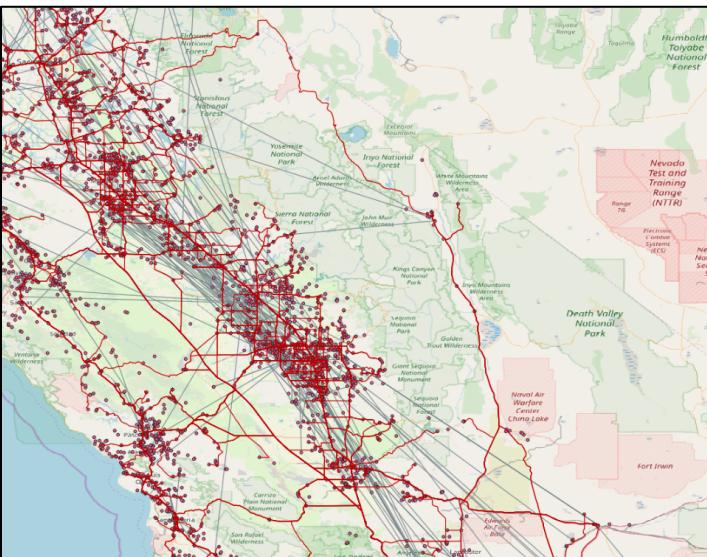


- Inference: Associations between attributes and activities in a network
- Prediction: Are there reproducible patterns that we can predict?

Graph theory

What is a graph? (in the context of network analysis)

“Mathematical representation of a network”



$$G = (V, E)$$

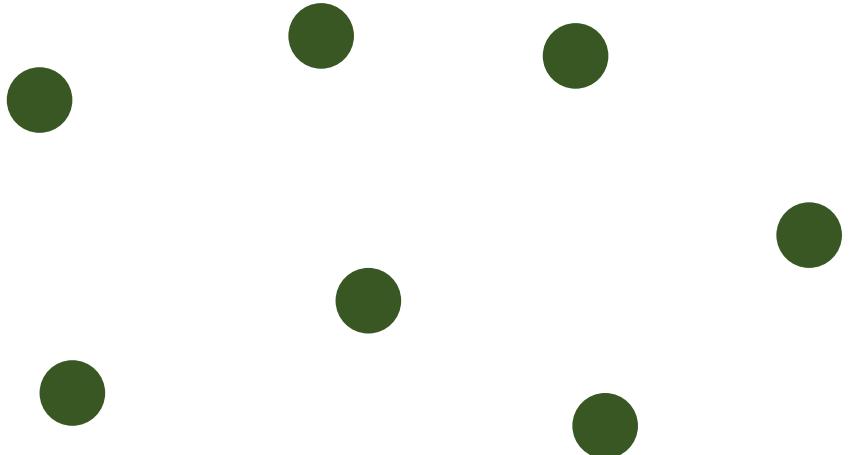
Where:

V = Vertice

E = Edge

I.B - Elements of a network

Nodes (vertices)



Agents or individuals forming a network

$$V = \{1, 2, 3, \dots, i\}$$

I.B - Elements of a network

Nodes (vertices)

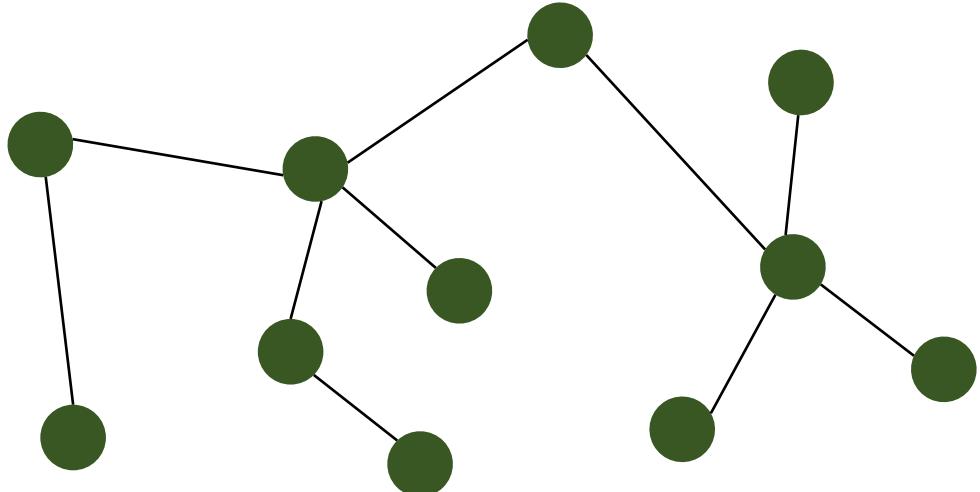


Agents or individuals forming a network:
farms, animals, humans, markets

$$V = \{1, 2, 3, \dots, i\}$$

I.B - Elements of a network

Edges (links)



Connection between a pair of nodes (dyad)

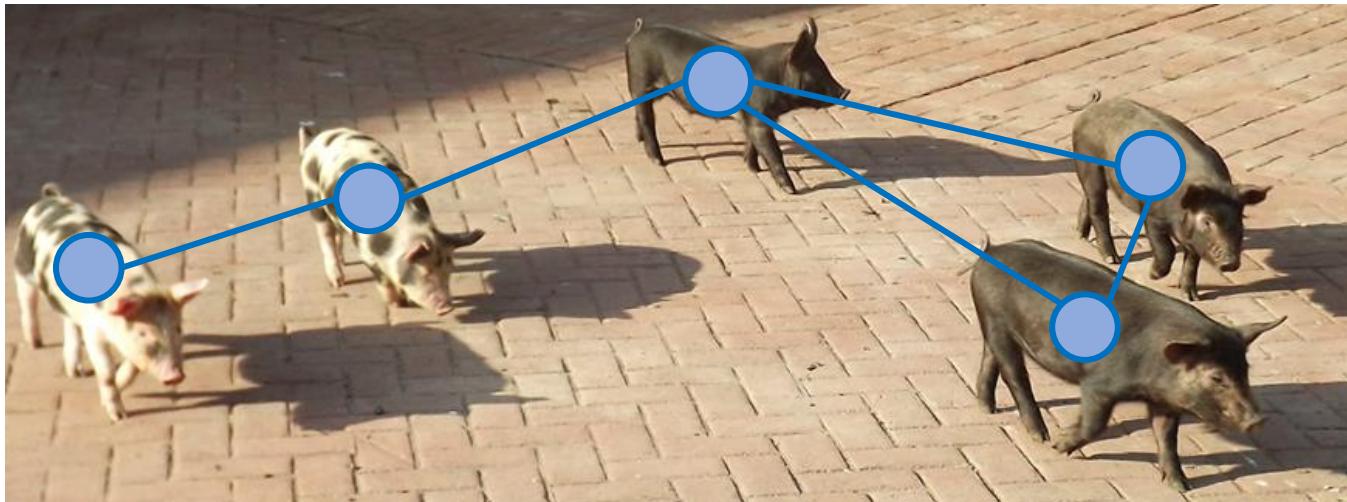
$$E = \{(1, 2), (1, 3), \dots, (i, j)\}$$

In a network, connected nodes are considered as **neighbours**

Each of the connected nodes belongs to a **neighborhood**

I.B - Elements of a network

Edges (links)



Connection between a pair of nodes (dyad):

Animal shipments, human movements, social contacts

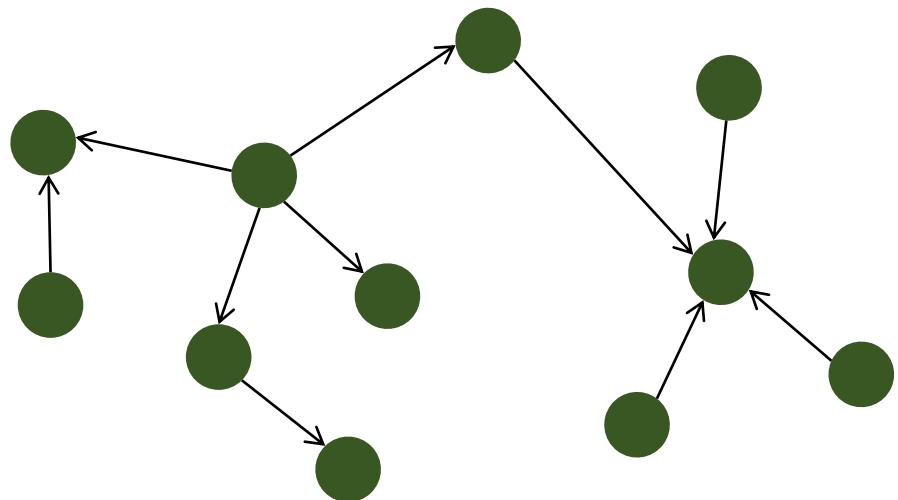
$$E = \{(1, 2), (1, 3), \dots, (i, j)\}$$

In a network, connected nodes are considered as **neighbours**

Each of the connected nodes belongs to a **neighborhood**

I.B - Elements of a network

Directionality



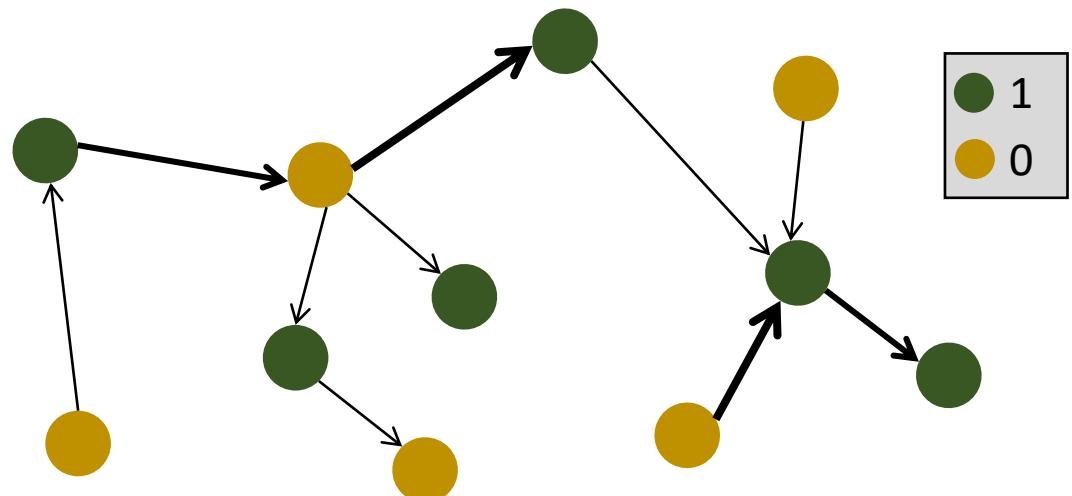
$$E = \{(1 \rightarrow 2), (1 \rightarrow 3), \dots, (i \rightarrow j)\}$$

I.B - Elements of a network

Atributes

Nodes: farm size, type...

Edges: movement size, cause...



$$V = \{0, 1, 0, \dots, i\}$$

$$E = \{(1), (3), \dots, (x_i)\}$$

I.C.a - Data sources

Define **nodes**:

- What is the unit of analysis (e.g. farm, animal, etc)

Define **edges**:

- Frequency of contacts
- Duration of contacts

I.C.a - Data structure

Nodes dataset			
ID	Farm	Farm size	Farm type
1	Swine and company	800	fattening farm
2	Les Cochonets	1200	sow farm
3	The farmhouse	50	small-scale farm
4	The Boar	20	reproductive males
5	Ham & sausage	65	slaughterhouse
6	The trading post	0	market

Edges dataset				
ID_origin	ID_destination	Shipment size	Date	Reason
2	6	10	02-25-2018	Sale of adult sows
2	1	15	03-12-2018	Piglets for fattening
4	1	1	10-15-2017	Insemination
1	5	100	06-30-2018	Slaughter
6	3	2	07-10-2018	Sale of adult sows
2	1	100	02-15-2018	Piglets for fattening

Measurement of contacts

- **Retrospective**

- National movement records
- Farm registry and population census
- Surveys

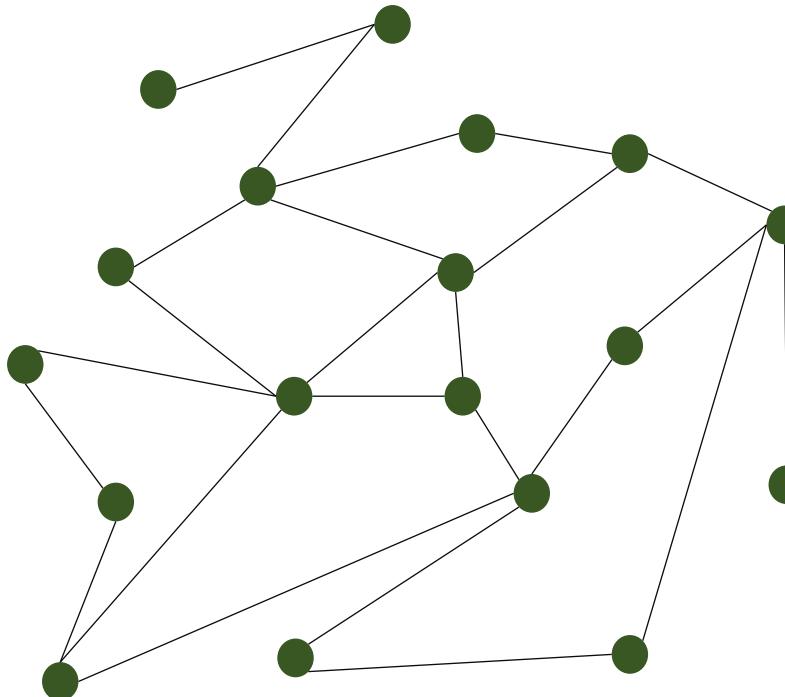
- **Prospective**

- Visual observation
- Tracking with GPS or PIT (Passive Integrated Transporter) systems

I.C.a - Data sources

Passive Surveillance

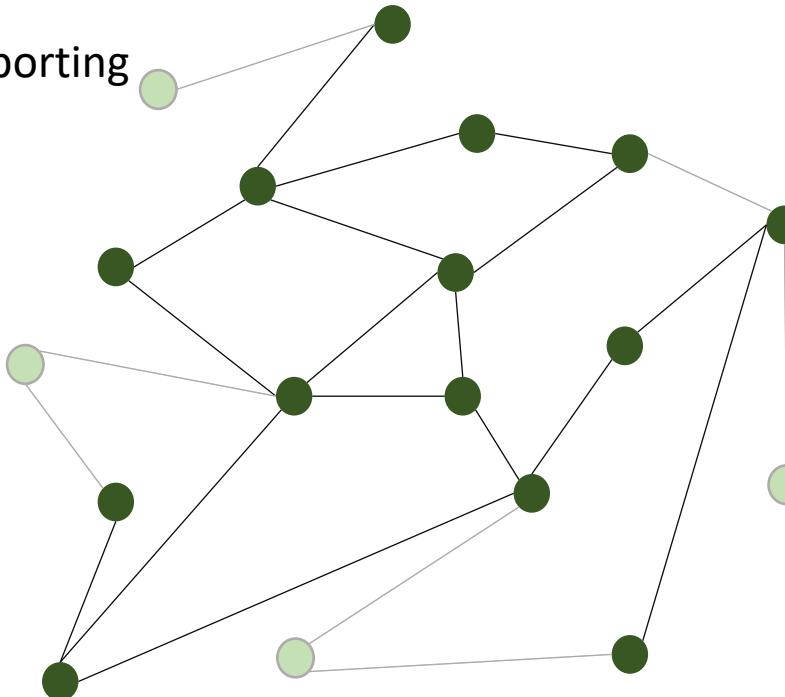
- Mandatory movement registries
- Population census
- GPS data (entire population)
 - Complete network



I.C.a - Data sources

Passive Surveillance

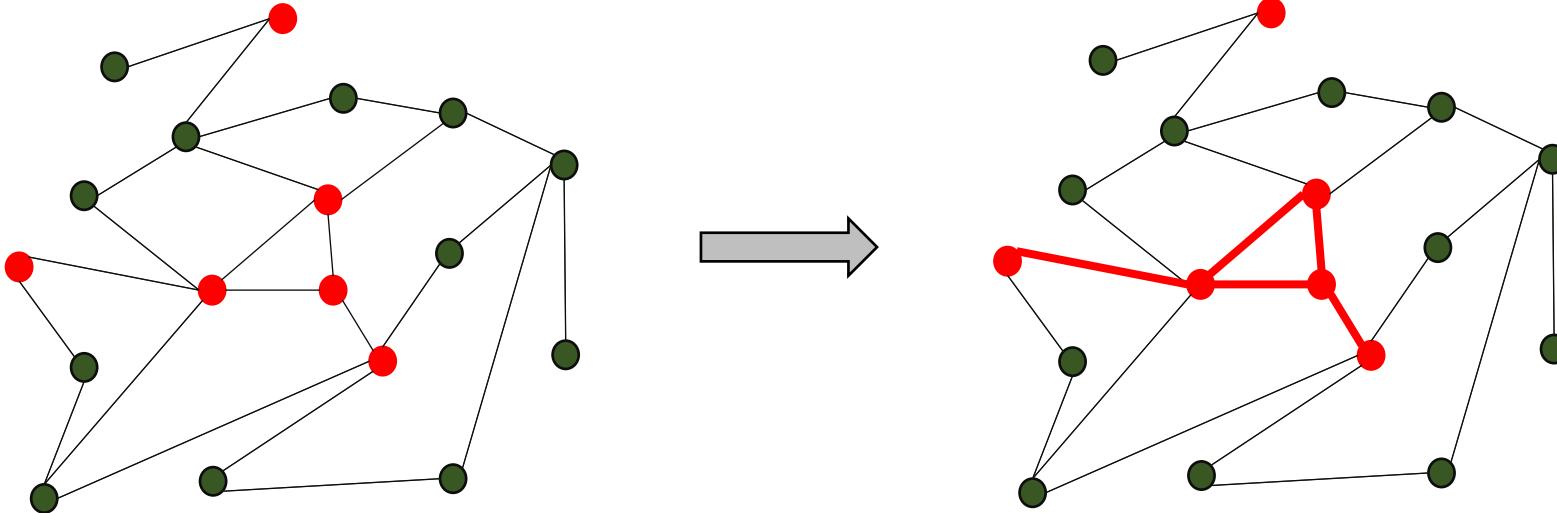
- Mandatory movement registries
- Population census
- GPS data (entire population)
 - Complete network
 - Impact of underreporting



I.C.a - Data sources

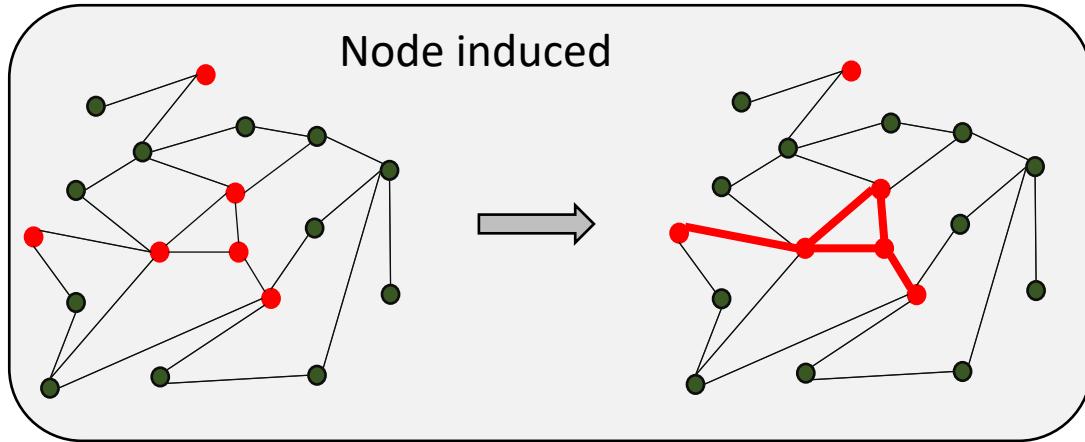
Active surveillance

- Surveys
- Observation of a sample of the population
 - Subgraph sampling and incomplete network

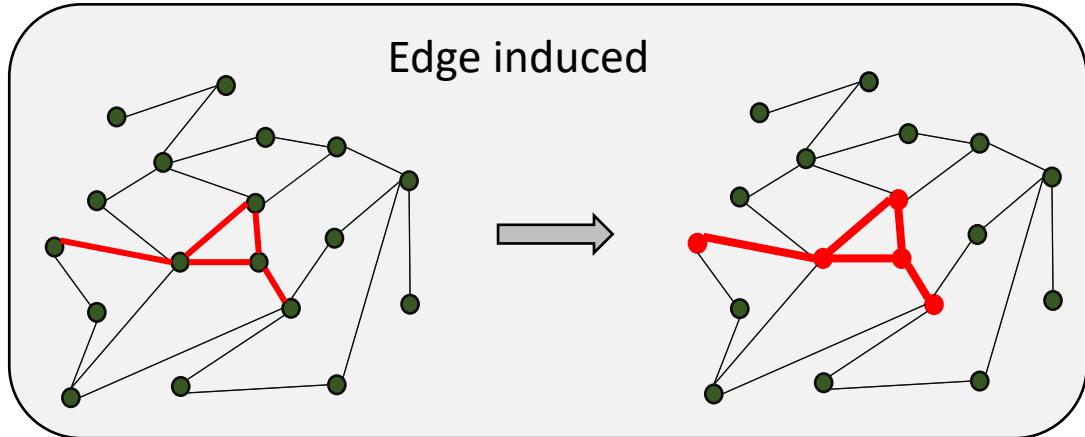


I.C.b - Sampling method

Random Sampling Methods



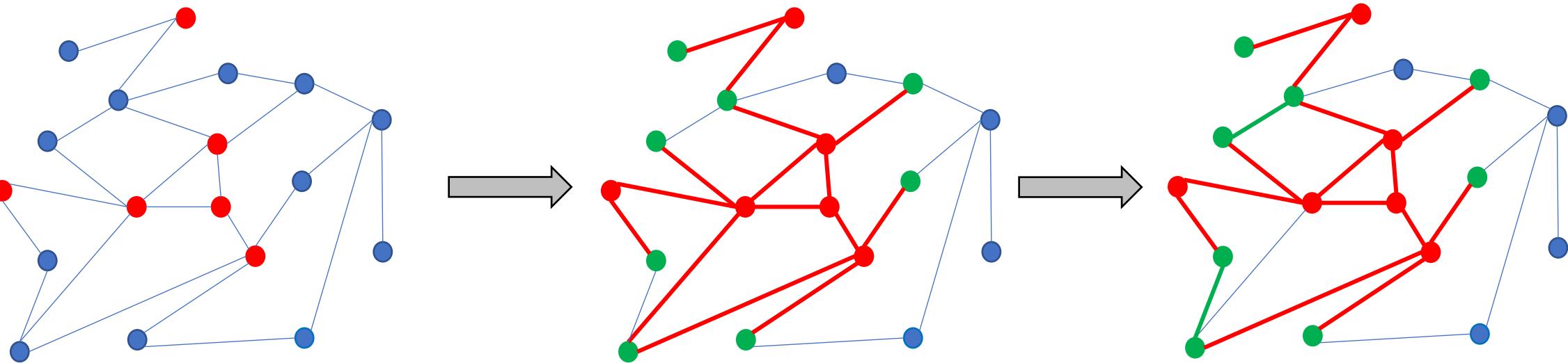
- Sample a group of nodes
- Identify all contacts between these nodes



- Sample a group of edges
- Identify all nodes connected to these edges

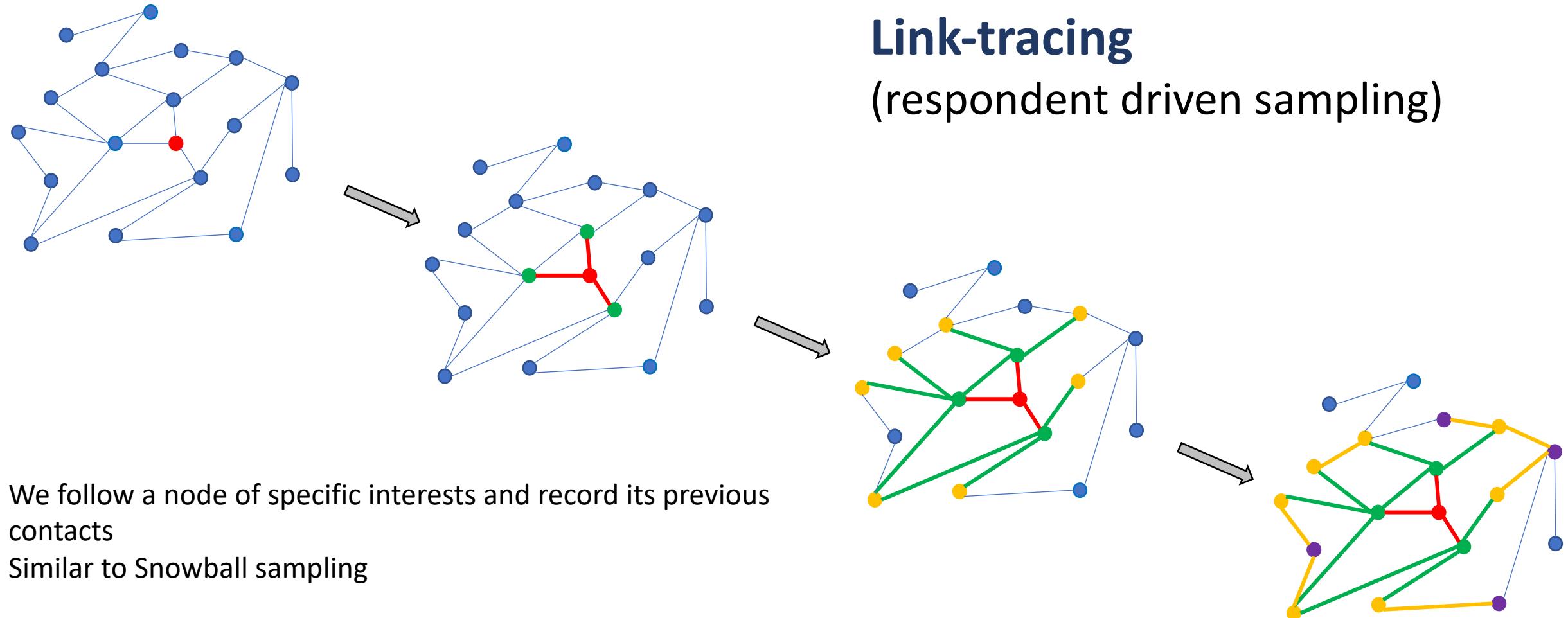
I.C.b - Sampling method

Egocentric Sampling



- Sample a group 1 of nodes
- Identify all contacts linked members of group 1
- Identify all nodes (group 2) directly connected to group 1
- Identify all contacts between members of group 2

I.C.b - Sampling method



Link-tracing
(respondent driven sampling)

Table of contents

I. Introduction to networks

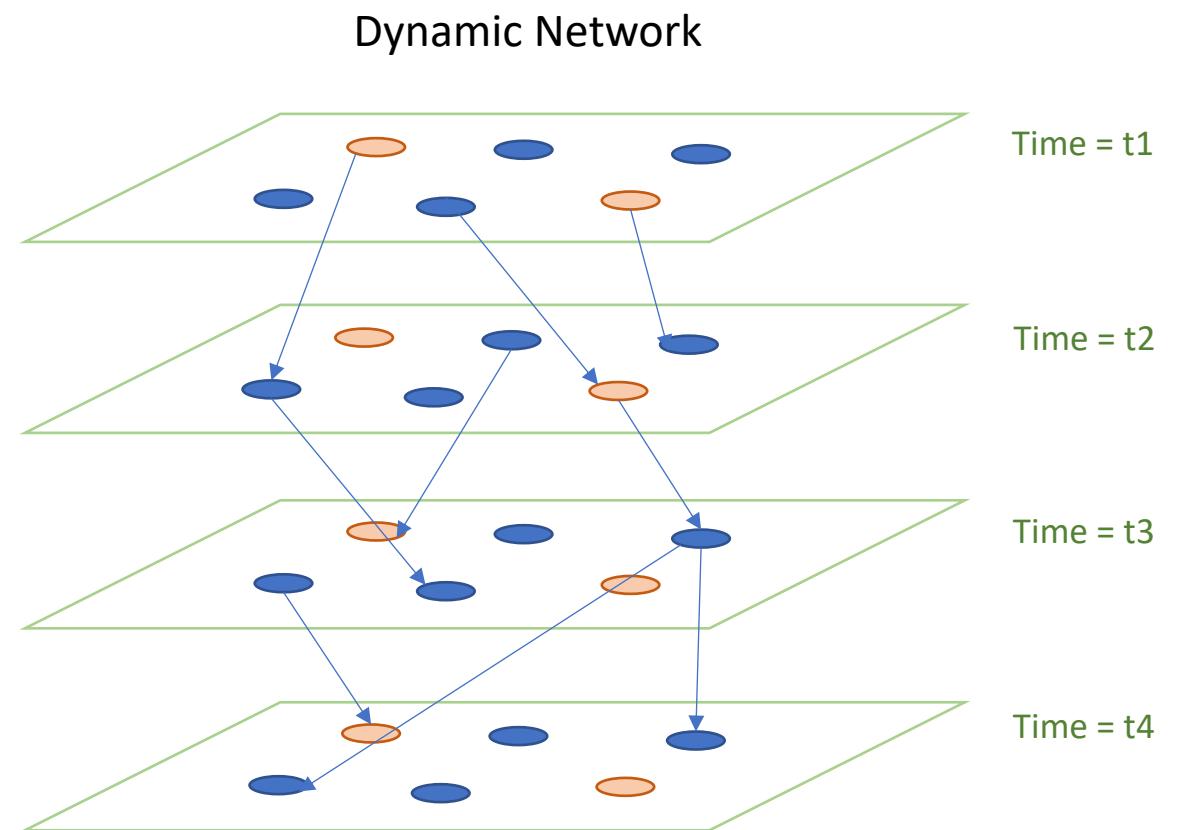
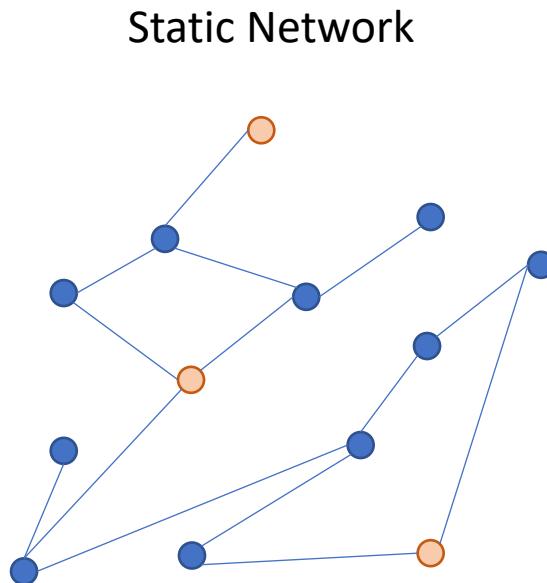
- A. What is a network?
- B. Elements of a network
- C. Data sources and Sampling methods

II. Static network analysis

- A. Network simplification
- B. Network properties
 - a) General properties
 - b) Local properties
- C. Weak and Strong component

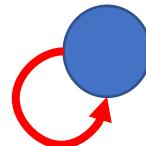
Static Network Analysis

- Network Type

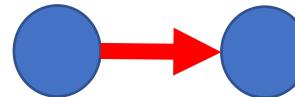
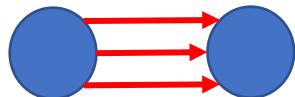


II.A - Network simplification

- Loops
 - Are they important?



- Multiple Edges
 - Meaning? Social/Commercial relationship? Useful for modeling?

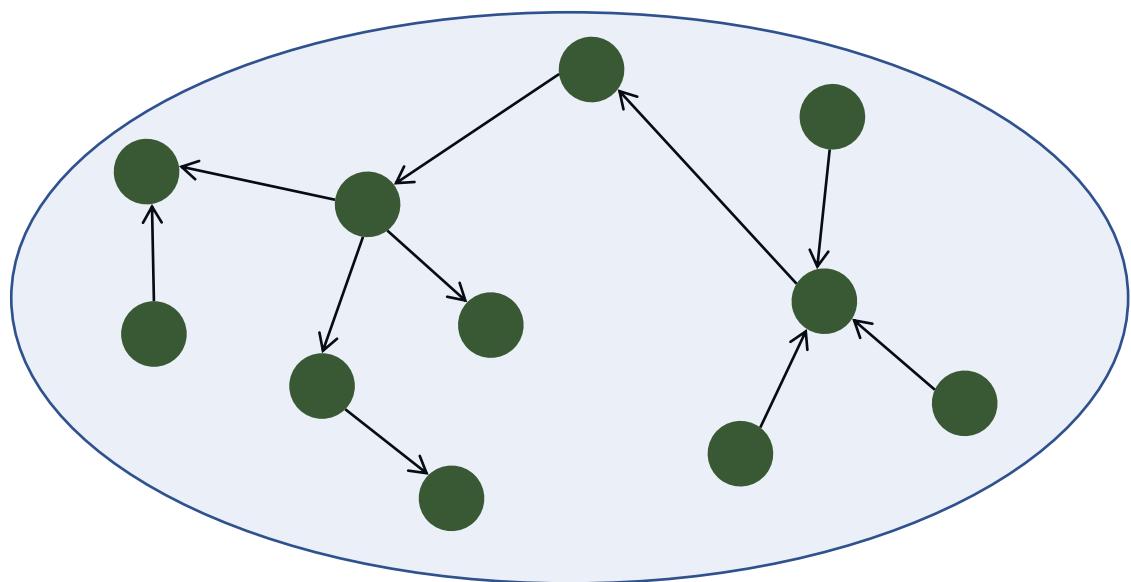


- Directed or undirected
 - Is it relevant to disease transmission in a given context: direct animal contact vs shipment

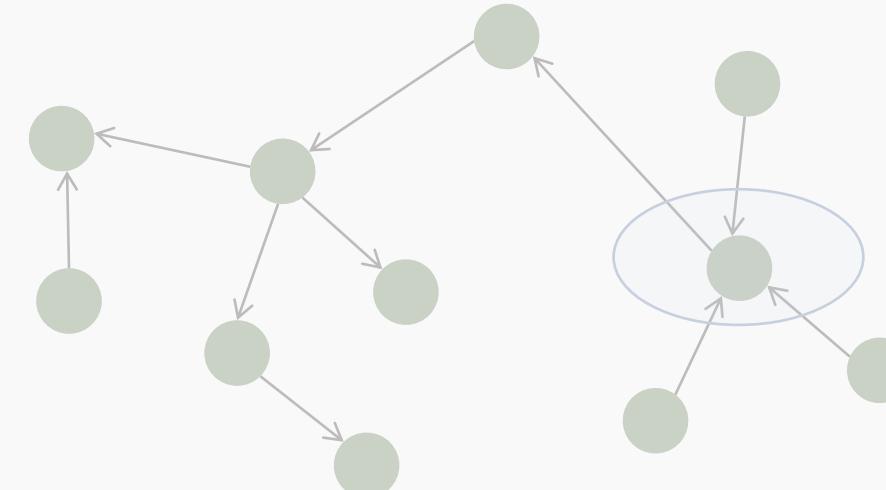


II.B - Network properties

General properties



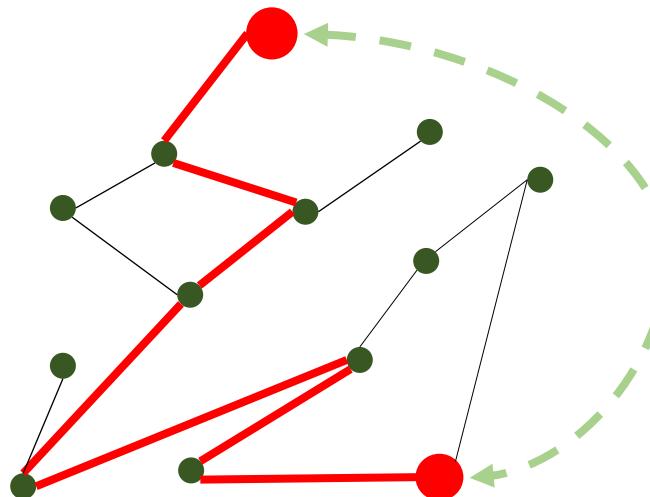
Local properties



II.B.a - General properties

Size: Number of nodes in the network (1 to ∞)

Diameter: The shortest path between the most distant nodes (1 to ∞)

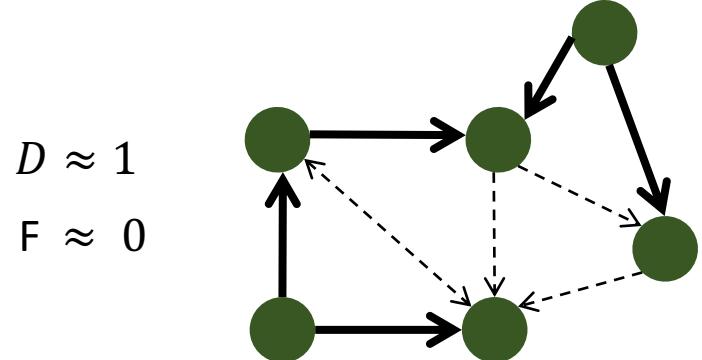


II.B.a - General properties

Density: Ratio of all observed edges in network to all theoretical edges that could exist (0 to 1)

$$D = \frac{E}{N(N - 1)}$$

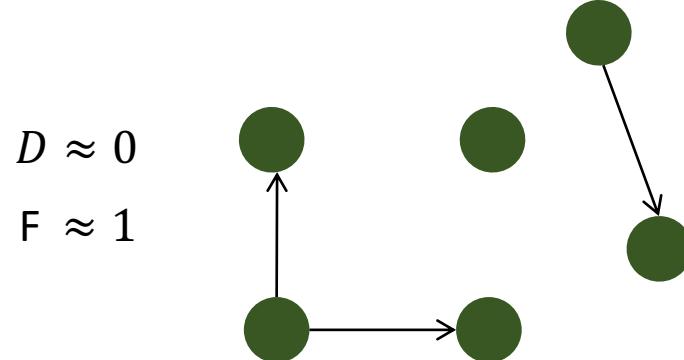
$$D = \frac{2E}{N(N - 1)}$$



$$D \approx 1$$
$$F \approx 0$$

- **Fragmentation**: Proportion of pairs of nodes that are not connected to each other. High fragmentation means more isolated nodes

$$F = 1 - \frac{2 \sum_i \sum_{j < i} r_{ij}}{N(N - 1)}$$



$$D \approx 0$$
$$F \approx 1$$

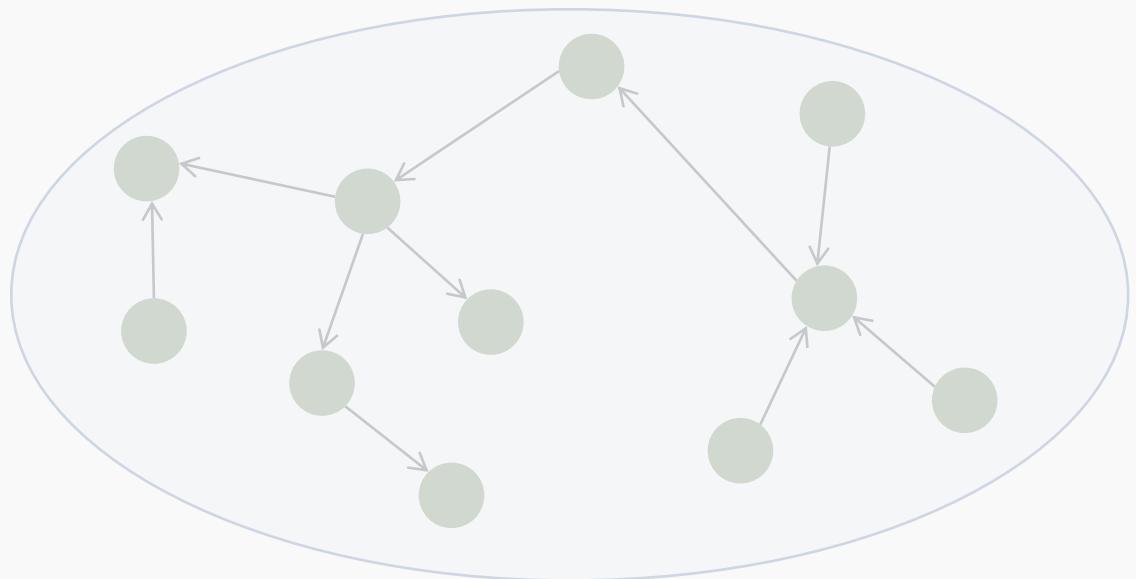
II.B.a - General properties

- **Clustering coefficient:** The probability of individual nodes being connected
- Values of 1 indicates that all nodes are directly connected to all others. Value of 0 indicates absence of contacts.

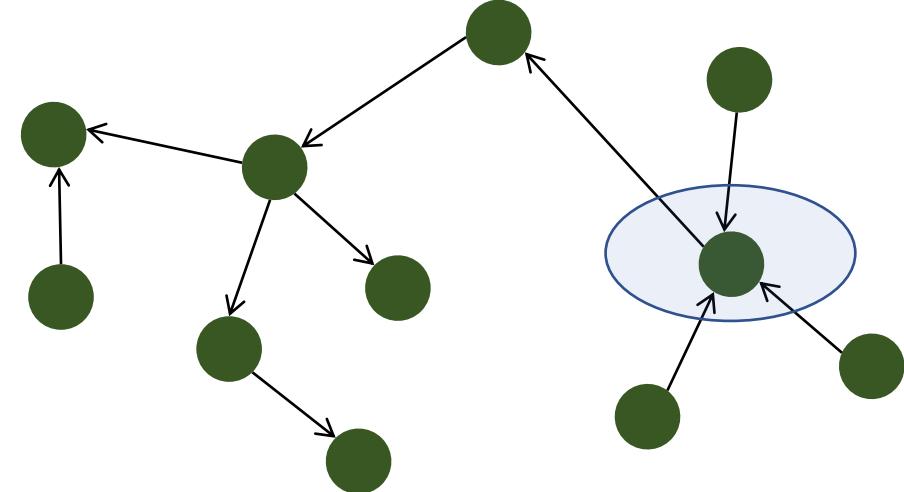
$$CC = \frac{1}{N} \sum_{i=1}^N \frac{e_{jz}}{k_i(k_i - 1)}$$

II.B - Network properties

General properties



Local properties



II.B.b - Local properties

Centrality measures:

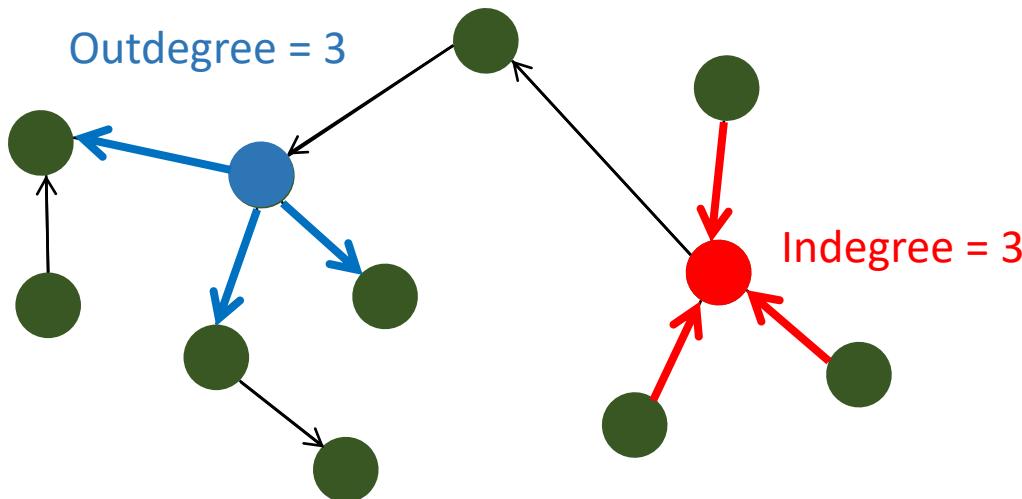
- Degree
- Betweenness
- Closeness

II.B.b - Local properties

Degree (Dc_i), Number of edges for each node.

Outdegree Do_i : Number of edges originating from node i .

Indegree Di_i : Number of contacts arriving at node i .



Relative degree centrality

$$RDc_i = \frac{Dc_i}{N - 1}$$

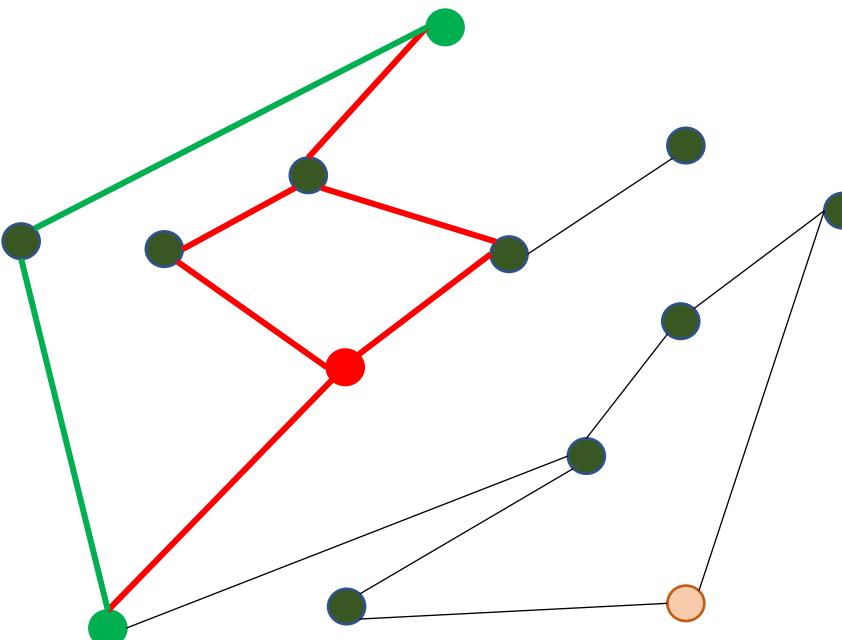
$$RDi_i = \frac{Di_i}{N - 1}$$

$$RDo_i = \frac{Do_i}{N - 1}$$

II.B.b - Local properties

Betwenness (Bc_i) Estimate of the probability that the shortest path between any pair of nodes passes through node i .

$$Bc_i = \sum \frac{\#SP \text{ through } i}{\#SP}$$



Relative betwenness centrality

For non-directed networks

$$RBc_i = \frac{Bc_i}{(N - 1)(N - 2)/2}$$

For directed networks

$$RBc_i = \frac{Bc_i}{(N - 1)(N - 2)}$$

II.B.b - Local properties

Closeness (Cc_i) Inverse of sum shortest paths between node i and all other nodes. Describes how close node i is to all other nodes.

Non-directed networks:

$$Cc_i = \frac{1}{\sum_{j=1}^N d_i(n_j, n_i)}$$

Relative closeness centrality

$$RCc_i = (N - 1) \times Cc_i$$

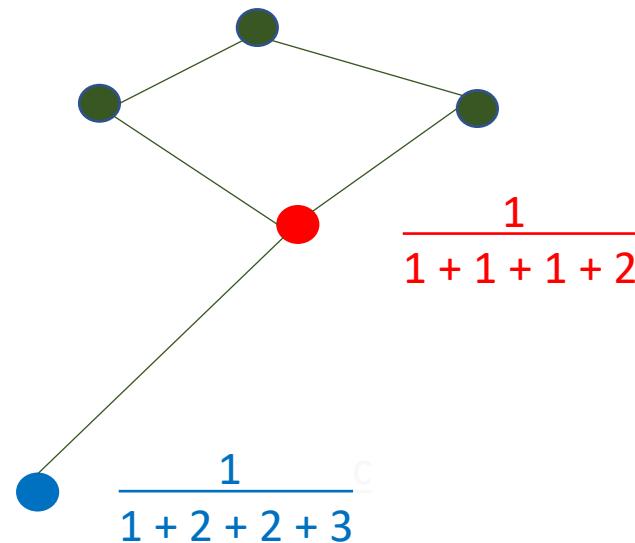
Directed networks:

In-closeness

$$Ci_i = \frac{1}{\sum_{j=1}^N d_i(n_j, n_i)}$$

Out-closeness

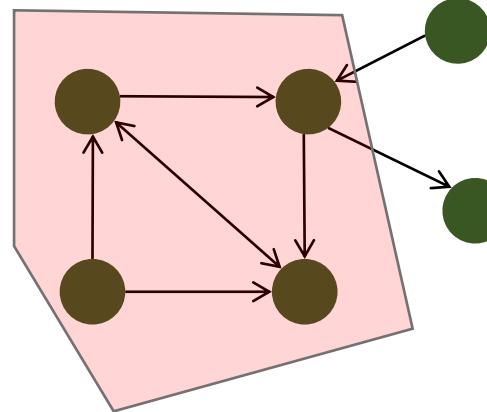
$$Co_i = \frac{1}{\sum_{j=1}^N d_o(n_j, n_i)}$$



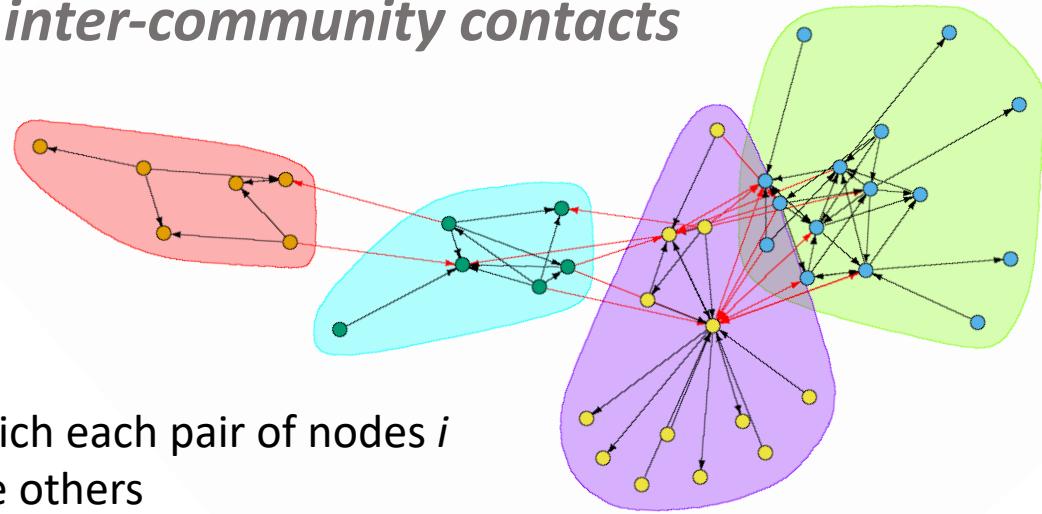
II.C - Detection of communities

Communities: Many intra-community contacts and few inter-community contacts

Strong Component

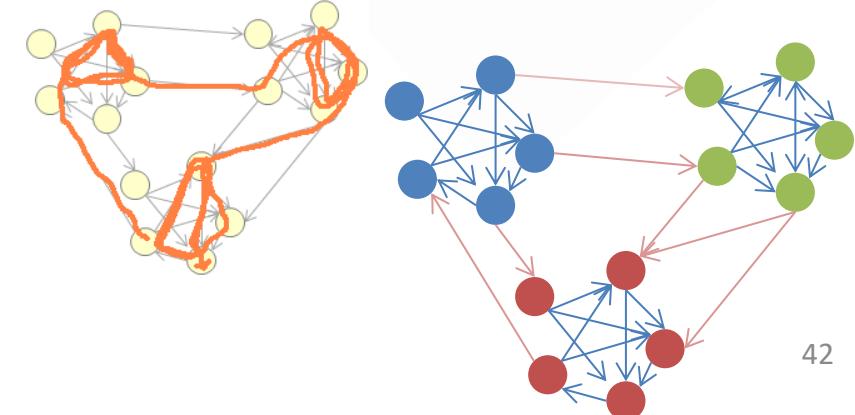


A subset of the network in which each pair of nodes i and j can be reached by all the others



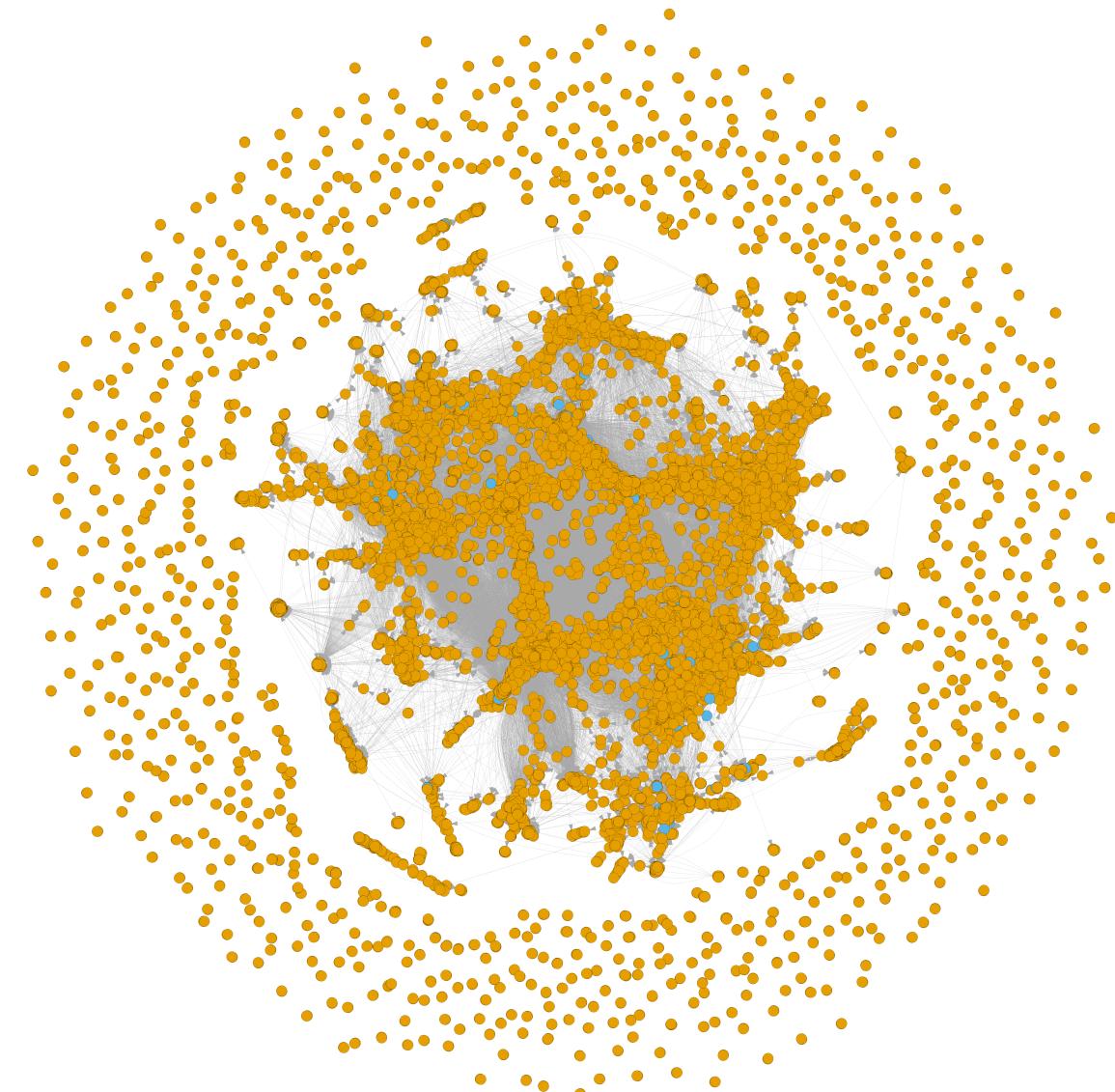
There are many algorithms to detect communities, **walktrap** being one of the most popular.

It generates random paths between nodes to detect zones where it circulates more frequently within a group of nodes.



Limitations of static network analysis and graph

- Network size and legibility
- No visible timeline in a static network
- Doesn't take into account spatial characteristics
 - Location based on network algorithm

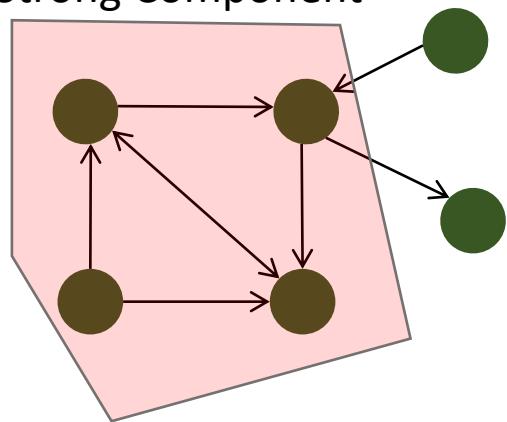


Questions?

Contact: jpgo@ucdavis.edu, jnbaron@ucdavis.edu
jpablo91.github.io

Questions?

Strong Component



Questions?

