

# Social Network Analysis from positioned telemetry data

Eneko Aspíllaga

Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB)  
Esportes, Balearic Islands, Spain

[aspillaga@imedea.uib-csic.es](mailto:aspillaga@imedea.uib-csic.es)  
<https://github.com/aspillaga>  
@enekoasp



CALCUATING POSITIONS IN AQUATIC TELEMETRY

ETN COST Action training school  
České Budějovice (Česká republika)

18-22 May 2022

# Session summary

## 1. Introduction to networks

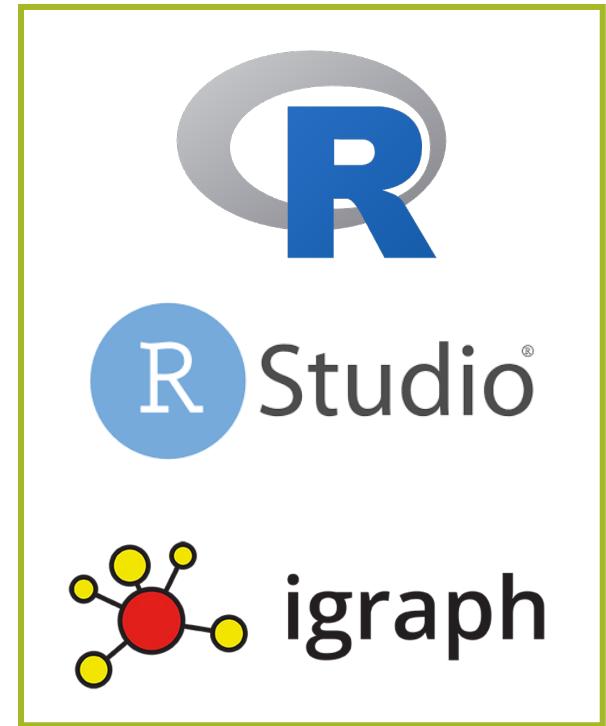
Practice: [SNA\\_01\\_introduction\\_to\\_networks.R](#)

## 2. Estimating associations from telemetry data

Practice: [SNA\\_02\\_network\\_from\\_tracks.R](#)

## 3. Null models for hypothesis testing

Practice: [SNA\\_03\\_null\\_models\\_hypothesis\\_testing.R](#)



# Main references

## Journal of Animal Ecology

*Journal of Animal Ecology* 2015, 84, 1144–1163

'HOW TO...' PAPER

### Constructing, conducting and interpreting animal social network analysis

Damien R. Farine<sup>1,2,3,\*</sup> and Hal Whitehead<sup>4</sup>



doi: 10.1111/1365-2656.12418

Farine and Whitehead 2015, J. Anim. Ecol.  
<https://doi.org/10.1111/1365-2656.12418>



## Methods in Ecology and Evolution

*Methods in Ecology and Evolution* 2017, 8, 1309–1320

doi: 10.1111/2041-210X.12772

### A guide to null models for animal social network analysis

Damien R. Farine<sup>\*1,2,3</sup>

Farine 2017, Meth. Ecol. Evol.

<https://doi.org/10.1111/2041-210X.12772>



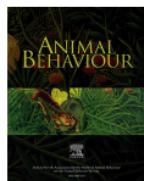
ELSEVIER

*Animal Behaviour* 104 (2015) e1–e5

Contents lists available at ScienceDirect

## Animal Behaviour

journal homepage: [www.elsevier.com/locate/anbehav](http://www.elsevier.com/locate/anbehav)



Forum

### Proximity as a proxy for interactions: issues of scale in social network analysis

Damien R. Farine <sup>a, b, c, \*</sup>



CrossMark

Farine 2015, Anim. Behav.

<https://dx.doi.org/10.1016/j.anbehav.2014.11.019>



Received: 12 January 2021 | Accepted: 1 October 2021

DOI: 10.1111/2041-210X.13741

## RESEARCH ARTICLE

### Permutation tests for hypothesis testing with animal social network data: Problems and potential solutions

Damien R. Farine<sup>1,2,3</sup> | Gerald G. Carter<sup>4,5</sup>

Farine and Carter 2021, Meth. Ecol. Evol.

<https://doi.org/10.1111/2041-210X.13741>

# 1. Introduction to networks

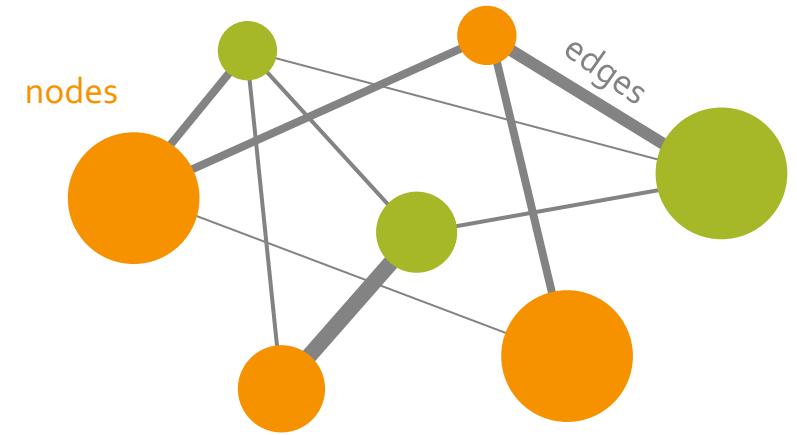


# Social Network Analysis (SNA)

- SNA is a tool to explore social structures based on interaction or association data
- Grounded on the graph theory
- Allows to address a broad set of biological and ecological questions:
  - Affiliation or avoidance behaviours, mating behaviour, connectivity, information flows, transmission of diseases, ...
  - Properties of individuals extracted from SNA can be linked to other biological traits (fitness, genetics, etc.)

# Network basics

- Observed pattern of interactions or associations
  - The underlying assumptions must be carefully considered for inherent biases (real vs observed networks)
- Represented as a graph
  - **Nodes or vertices:**
    - Individuals, groups, species, ...
    - Each node might have some specific attributes
  - **Links or edges:**
    - Define how nodes relate to each other
    - Binary (0 or 1, presence or absence of relationship) or weighted (numeric values indicating the strength of relationships)
    - Undirected vs directed



Undirected



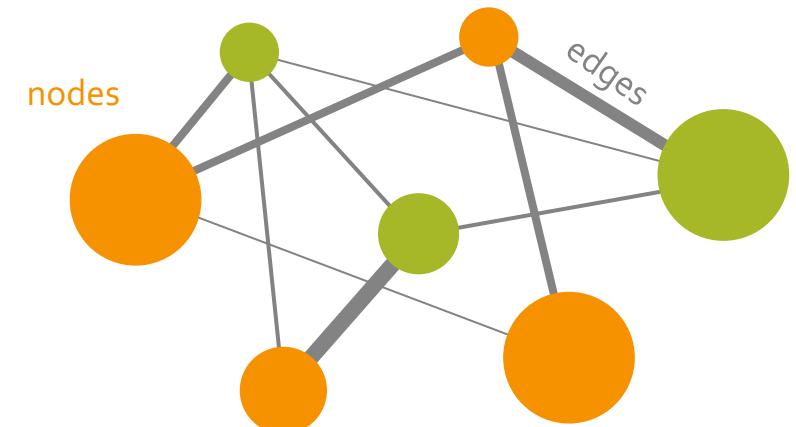
Directed



# Representation of networks

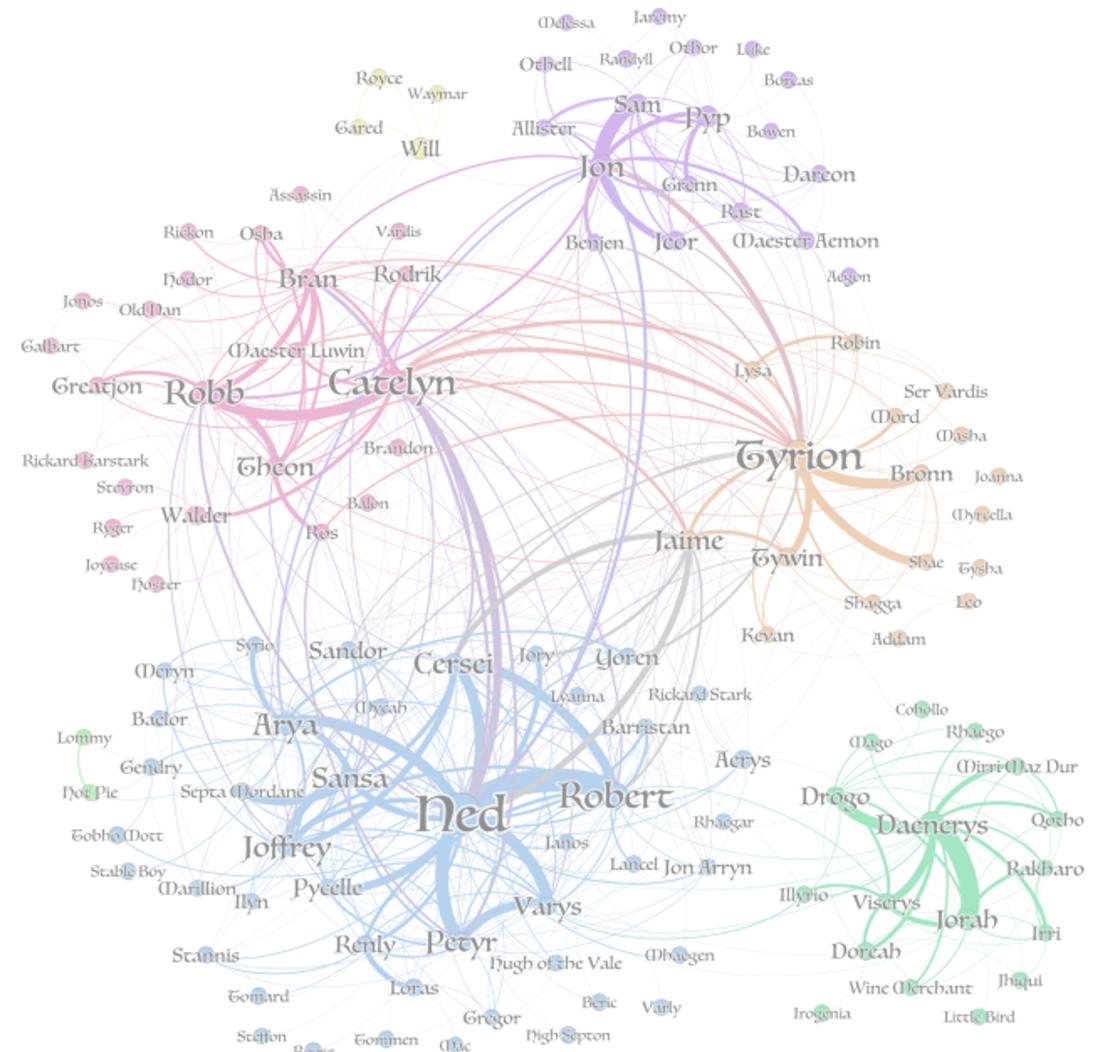
- Usually represented as an adjacency matrix
  - Rows and columns represent nodes ( $n \times n$ )
  - Values of the matrix represent the weight of each edge
  - Diagonals represent self-edges (rarely used in animal SNA, so usually equal to 0)
  - Symmetrical (undirected networks) or asymmetrical (directed networks)
- Visualized as network diagrams
  - 2D representations of the network
  - Nodes arranged depending on their connections within the graph
  - Node and edge attributes or network metrics can be represented

	ID 1	ID 2	ID 3	ID 4	...	ID n
ID 1	0	0.2	0	0.4	...	0.8
ID 2	0.2	0	1	0.7	...	0.2
ID 3	0	1	0	0	...	0.3
ID 4	0.4	0.7	0	0	...	0
...	...	...	...	...	...	...
ID n	0.8	0.2	0.3	0	...	0



# Creating our first network in R

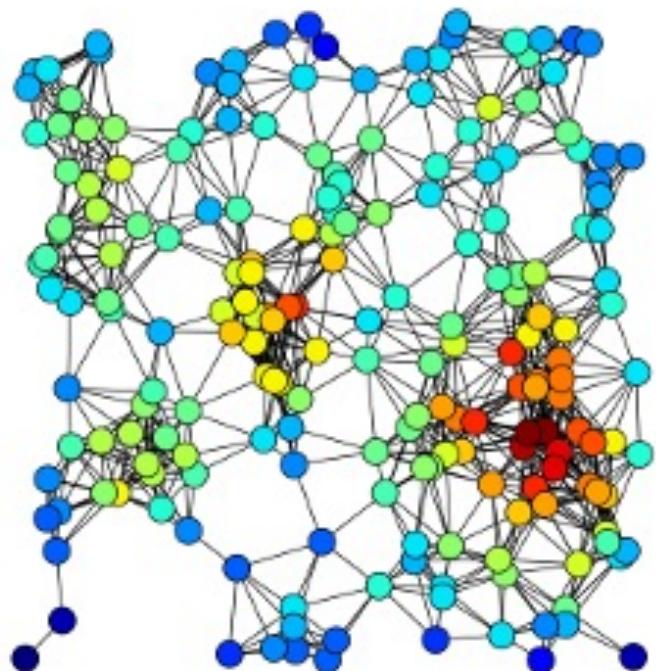
- Script [SNA\\_01\\_introduction\\_to\\_networks.R](#)
- Example dataset: Interactions between characters in the Season 1 of **Game of Thrones**, HBO series (Beveridge and Chemers, 2018)
  - URL: <https://networkofthrones.wordpress.com/>



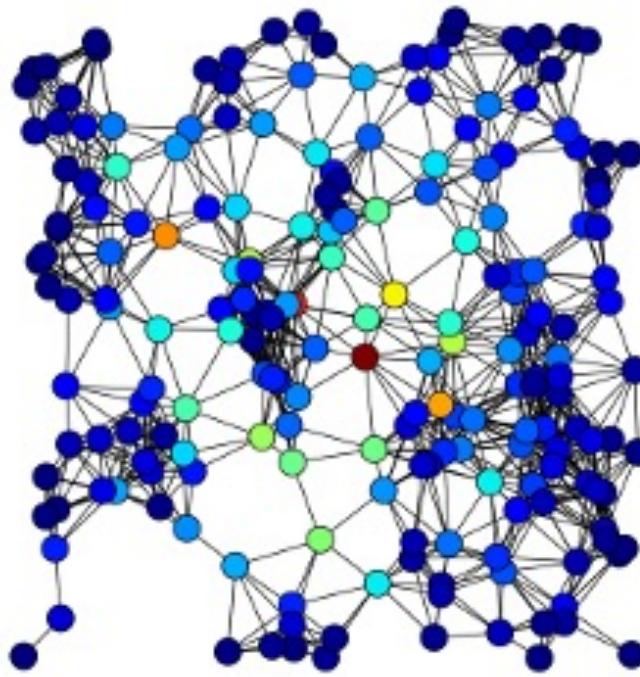
# Network metrics

- Parameters of the network can be used to describe its structure and to characterize the importance of each node connecting the network
- Network-level metrics:
  - **Number of edges**
  - **Edge-weight distribution**
  - **Edge density:** number of observed edges / number of possible edges
- Node-level metrics:
  - **Degree (binary degree):** number of edges of a node. Indicates the number of associates or interaction partners.
  - **Strength (weighted degree):** sum of all the edge weights of a node. Represents the association rate per node.
  - **Betweenness centrality:** count of the number of the shortest paths going through a node. Indicates how important a node is in connecting different parts of the network
  - **Eigenvector centrality:** sum of the centralities of the neighbours of a node. Indicates the influence of a node in a network. High centralities are reached by having a high degree or by being connected to associates with a high degree

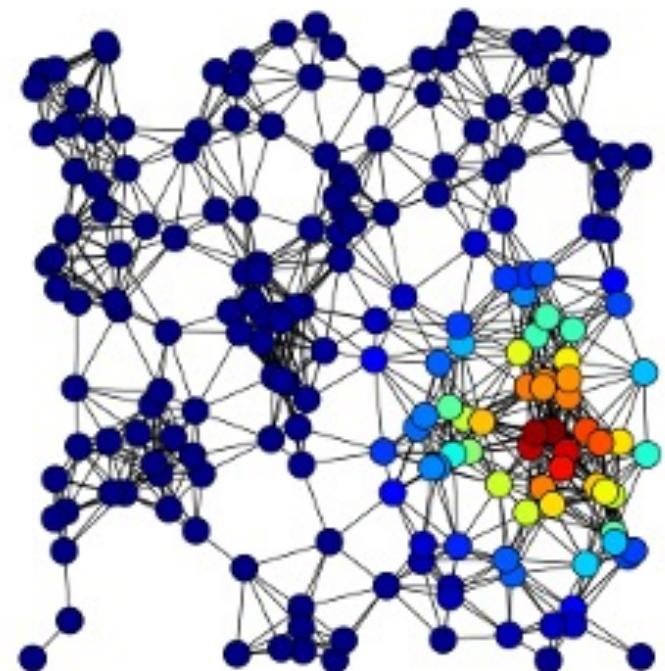
## Differences between node-metrics (centrality measurements)



Degree



Betweenness centrality

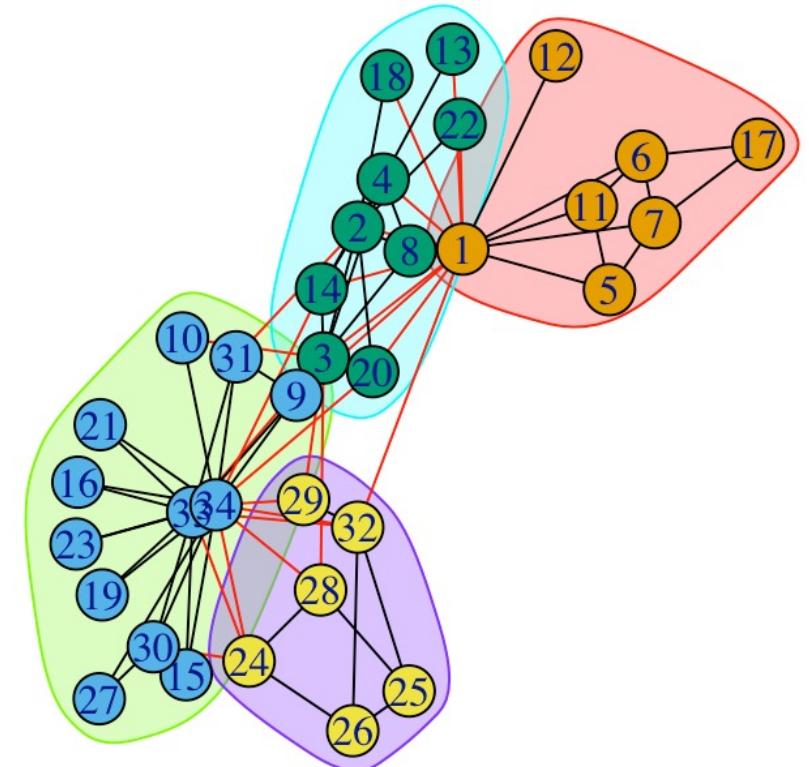


Eigenvector centrality

Source: <https://en.wikipedia.org/wiki/Centrality>

# Identifying communities

- Some networks are highly modular, with nodes forming densely connected communities
- Many techniques to cluster nodes into communities are already implemented in R

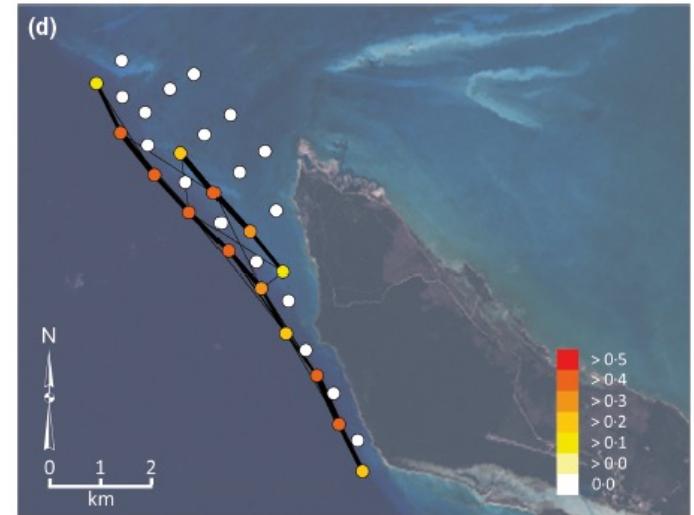
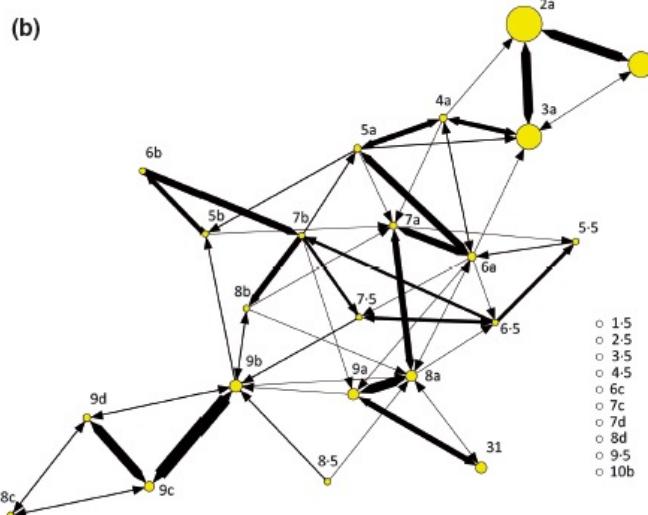
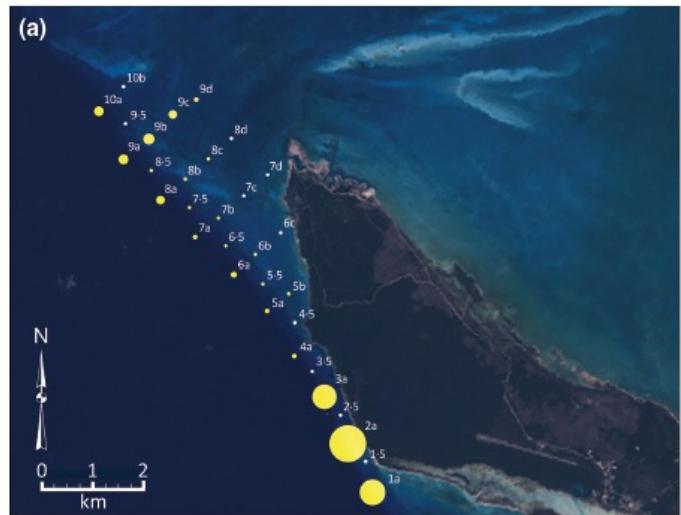


## 2. Estimating associations from telemetry data



# Network analysis with conventional telemetry data

- The network analysis approach has been widely used to analyse conventional telemetry data (presence/absence data in discrete receiver locations)
  - Spatial networks (receivers as nodes)
  - Social networks (individuals as nodes)



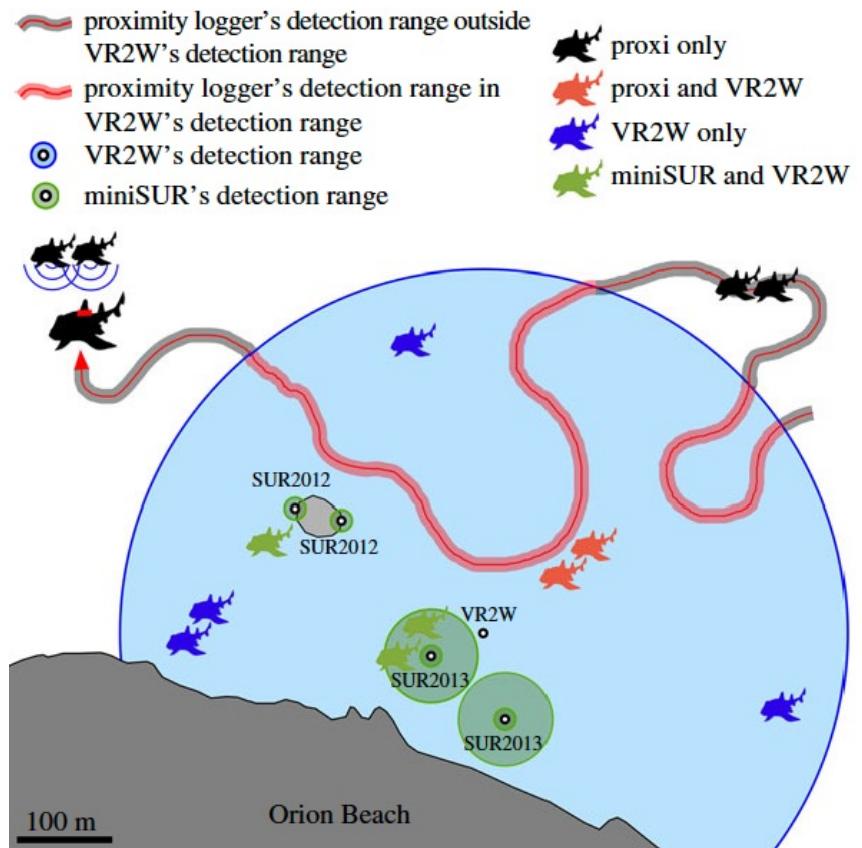
**Spatial networks**

Connectivity measures / Changes in space use

Source: Jacoby *et al.* 2012, MEE  
<https://doi.org/10.1111/j.2041-210X.2012.00187.x>

# SNA from conventional telemetry data

- Advantages of acoustic telemetry data:
  - Studied individuals are always unequivocally identified by their transmitter ID
  - Large numbers of individuals can be monitored simultaneously
- But measuring interactions is highly problematic
  - Large detection ranges (hundreds of meters) involve a large positioning uncertainty
    - Are individuals detected by the same receiver interacting between them?
    - Differences in detection ranges generate different association measurements (Mourier et al. 2017).
  - Not all the individuals in the population are monitored (most of interactions are missed)



Source: Mourier *et al.* 2017, R. Soc. Open Sci.  
<http://dx.doi.org/10.1098/rsos.170485>

# SNA from positioned telemetry data

- Positioned detections (trajectories) reduce the positioning uncertainty and help making certain assumptions
- Proximity-based networks:
  - Been close to another is a requisite for most interactions
  - Proximity networks might differ from interaction networks (depends on the scale of the studied subject – individuals, groups, populations)
- Associations (vs interactions):
  - Defined as the spatial and temporal co-occurrence of pairs of individuals
  - More adequate when interactions are rare or difficult to observe (as underwater)
  - Can capture a broad range of interactions that better represent the social environment
  - May provide more precise estimates of interaction rates by avoiding false negatives (interactions that occurred but were not observed)

# Estimating association indices

- Deciding how social ties are defined is a major step when applying SNA
  - Edge definition should be relevant to the question or hypothesis being addressed and our ability to observe the study system
  - Associations should not be defined by arbitrary thresholds (e.g. distance thresholds) that are not clearly defined by the study system
- **Association indices** usually indicate the proportion of time in which individuals are associated (0 - 1)

## Simple ratio index

Used when observations are rarely missed

$$E_{AB} = \frac{x}{x + y_{AB} + y_A + y_B}$$

## Half-weight index

Used if observations are frequently missed

$$E_{AB} = \frac{x}{x + y_{AB} + \frac{y_A + y_B}{2}}$$

## Both identified

Controls for cooccurrence

$$E_{AB} = \frac{x}{x + y_{AB}}$$

$E_{AB}$  = Association between individuals A and B

$x$  = Observation in which A and B were associated

$y_{AB}$  = Observation in which A and B were observed not associated

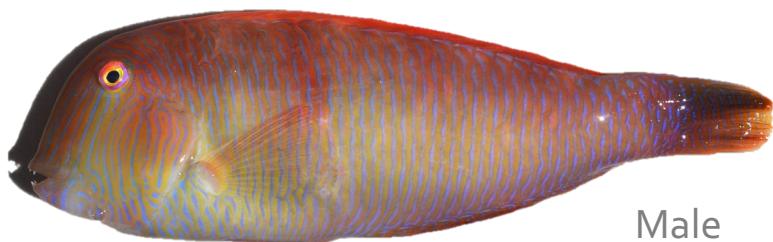
$y_A, y_B$  = Observations in which only A or B was observed

Whitehead 2008, *Analyzing Animal Societies*

<http://dx.doi.org/10.7208/chicago/g780226895246.001.0001>

# Creating a proximity network from trajectory data

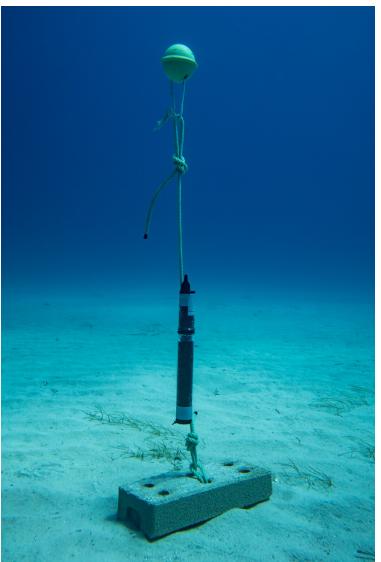
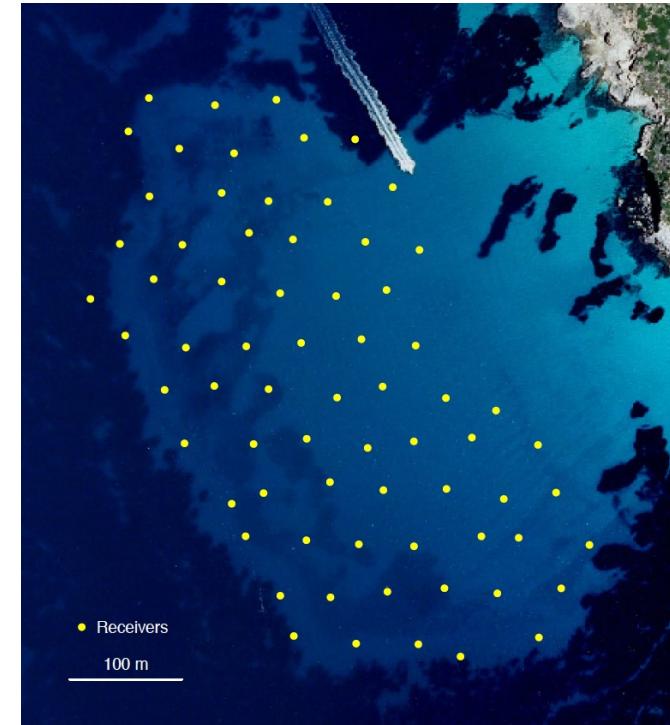
- Script [SNA\\_02\\_network\\_from\\_tracks.R](#)
- Example dataset: High-resolution trajectories of pearly razorfish (*Xyrichtys novacula*) individuals obtained using the Lotek JSATS telemetry system in the Palma Bay marine reserve (Mallorca, Spain)
  - Information about the telemetry system and the positioning and post-processing workflow can be found in Aspíllaga *et al.* 2021a (Anim. Biote.) <https://doi.org/10.1186/s40317-020-00224-w>
  - A more detailed study on the social interactions of the studied population can be found in Aspíllaga *et al.* 2022 (Front. Mar. Sci.) <https://doi.org/10.3389/fmars.2021.688010>



Male



Female



### 3. Null models for hypothesis testing



# Hypothesis testing

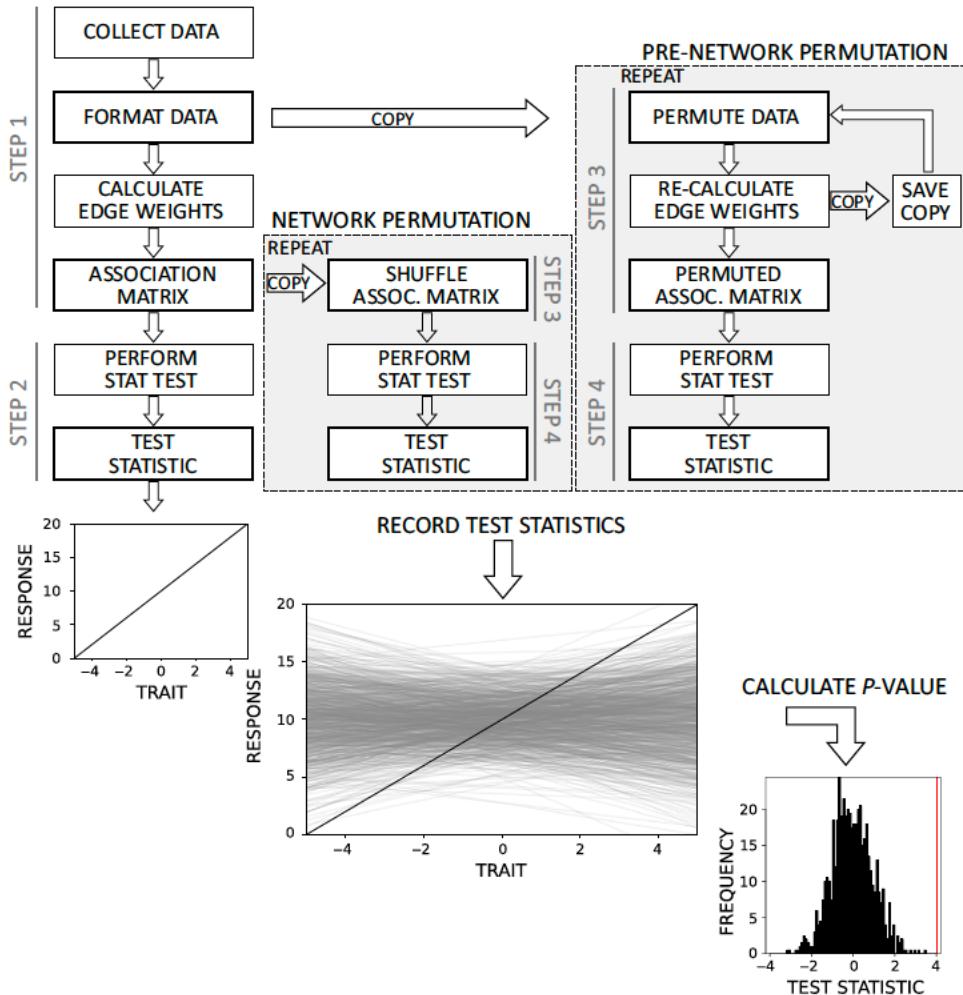
- Network measurements are inherently non-independent, violating most assumptions underlying most parametrical tests
- **Null-model:** data set based on the observed data but generated in a way that allows some aspects to be random, breaking the assumptions of the association patterns
  - Usually consists on shuffling the data to create expectations of random given certain constraints
  - But it might also involve generating new simulated data
- Observed networks are then compared against many iterations of null models

## Steps:

1. Generate a network from the observed data (**observed network**)
2. Decide and calculate the **test statistic** with the observed network
  - o If a model is applied to node metrics, the slope coefficients from linear or mixed models can be used
  - o To obtain the significance of network level metrics, the metric itself can be used (e.g. the mean degree, the coefficient of variation of edge weights)
3. Randomize the observed data and generate **null networks**
4. Calculate the test statistic using the same method but with the null networks
5. Compare the distribution of the test statistic in the null networks against the observed value to obtain a “P-value”

# Steps:

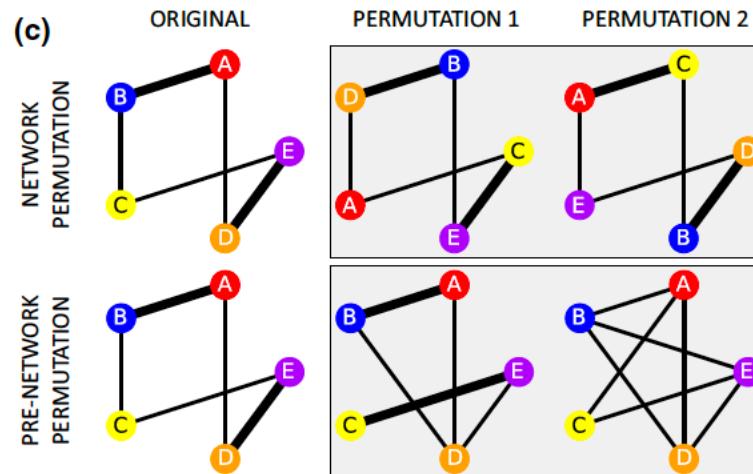
(a)



(b)

NETWORK PERMUTATION			PRE-NETWORK PERMUTATION		
ORIGINAL	SHUFFLE 1	SHUFFLE 2	ORIGINAL	SWAP 1	SWAP 2
ID	TIME	ID	TIME	ID	TIME
A	1	B	1	C	1
A	2	B	2	C	2
A	3	B	3	C	3
B	2	D	2	A	2
B	3	D	3	A	3
B	4	D	4	A	4
C	4	[A]	4	E	4
C	5	[A]	5	E	5
C	6	[A]	6	E	6
D	1	E	1	B	1
D	7	E	7	B	7
D	8	E	8	B	8
E	6	C	6	D	6
E	7	C	7	D	7
E	8	C	8	D	8

(c)

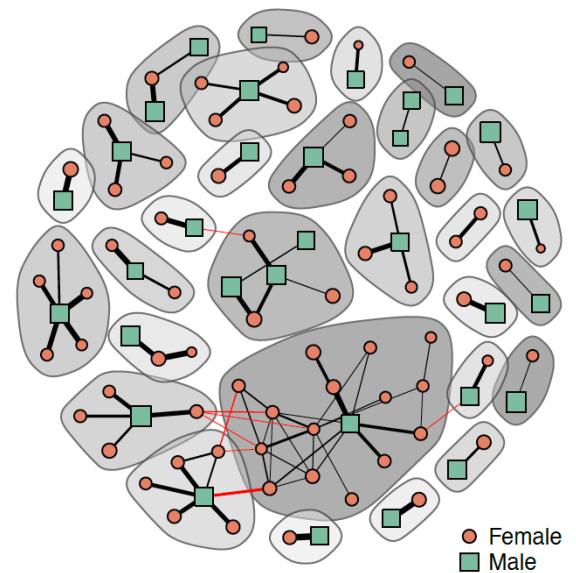
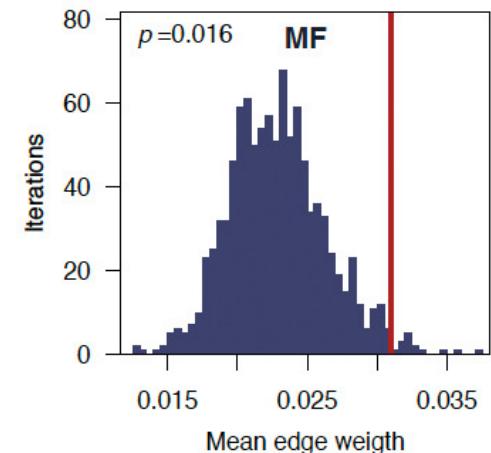
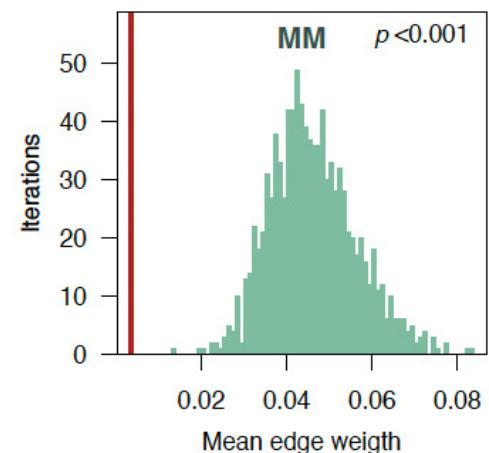
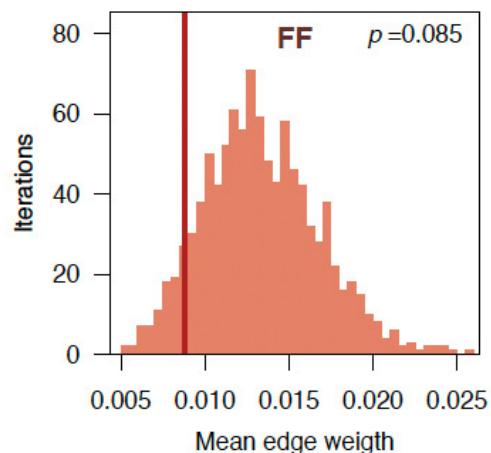


# Generating null models

- **Network permutations:** randomization is made after the network is generated
  - **Node-based randomizations (node permutation test):** the identity of each node is randomized, breaking the relation between the network and the individual traits
    - The attributes are entirely redistributed among the nodes, keeping the same number of each
    - Usually used to test the effect of node attributes in their position within the network
    - Simple to implement
    - Potential of high type I and type II errors
- **Pre-network permutations:** randomization is applied on the original data, before calculating the association indices
  - Enable a better control over the model and the hypothesis tested
  - Example (applied to trajectories): changing the date of the tracks of different individuals, randomly distributing trajectories
  - Preferred over network permutations (reduce both type I and type II errors)
  - Sometimes difficult to define

# Testing the significance of the razorfish social network

- Script [SNA\\_03\\_null\\_models\\_hypothesis\\_testing.R](#)
- We will continue with the analysis of the previous script, applying a pre-network randomization approach to test the significance of the estimated network parameters
  - More details on this study can be found in Aspíllaga *et al.* 2022 (Front. Mar. Sci.) <https://doi.org/10.3389/fmars.2021.688010>



# Bibliography

Aspíllaga, E., Arlinghaus, R., Martorell-Barceló, M., Barcelo-Serra, M., & Alós, J. (2021). High-Throughput Tracking of Social Networks in Marine Fish Populations. *Frontiers in Marine Science*, 8, 1–12. <https://doi.org/10.3389/fmars.2021.688010>

Aspíllaga, E., Arlinghaus, R., Martorell-Barceló, M., Follana-Berná, G., Lana, A., Campos-Candela, A., & Alós, J. (2021). Performance of a novel system for high-resolution tracking of marine fish societies. *Animal Biotelemetry*, 9(1), 1–14. <https://doi.org/10.1186/s40317-020-00224-w>

Farine, D. R. (2015). Proximity as a proxy for interactions: Issues of scale in social network analysis. *Animal Behaviour*, 104, e1–e5. <https://doi.org/10.1016/j.anbehav.2014.11.019>

Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, 8(10), 1309–1320. <https://doi.org/10.1111/210X.12772>

Farine, D. R., & Carter, G. G. (2022). Permutation tests for hypothesis testing with animal social network data: Problems and potential solutions. *Methods in Ecology and Evolution*, 13(1), 144–156. <https://doi.org/10.1111/210X.13741>

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *Journal of Animal Ecology*, 84(5), 1144–1163. <https://doi.org/10.1111/1365-2656.12418>

Hoppitt, W. J. E., & Farine, D. R. (2018). Association indices for quantifying social relationships: how to deal with missing observations of individuals or groups. *Animal Behaviour*, 136, 227–238. <https://doi.org/10.1016/j.anbehav.2017.08.029>

Jacoby, D. M. P., Brooks, E. J., Croft, D. P., & Sims, D. W. (2012). Developing a deeper understanding of animal movements and spatial dynamics through novel application of network analyses. *Methods in Ecology and Evolution*, 3(3), 574–583. <https://doi.org/10.1111/j.2041-210X.2012.00187.x>

Jacoby, D. M. P., & Freeman, R. (2016). Emerging Network-Based Tools in Movement Ecology. In *Trends in Ecology and Evolution*, 31(4), 301–314. <https://doi.org/10.1016/j.tree.2016.01.011>

Mourier, J., Bass, N. C., Guttridge, T. L., Day, J., & Brown, C. (2017). Does detection range matter for inferring social networks in a benthic shark using acoustic telemetry? *Royal Society Open Science*, 4(9), 170485. <https://doi.org/10.1098/rsos.170485>

Spiegel, O., Sih, A., Leu, S. T., & Bull, C. M. (2018). Where should we meet? Mapping social network interactions of sleepy lizards shows sex-dependent social network structure. *Animal Behaviour*, 136, 207–215. <https://doi.org/10.1016/j.anbehav.2017.11.001>

Villegas-Ríos, D., Jacoby, D. M. P., & Mourier, J. (2022). Social networks and the conservation of fish. *Communications Biology*, 5(1), 178. <https://doi.org/10.1038/s42003-022-03138-w>