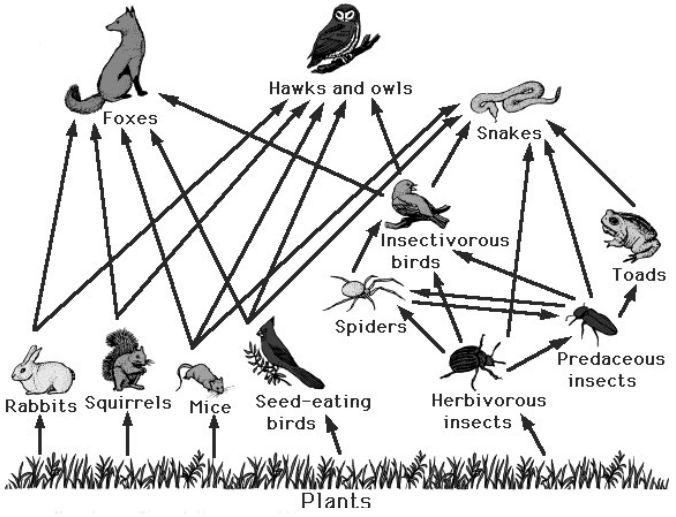
Model-based clustering of a collection of networks

Tabea Rebafka

LPSM, Sorbonne Université
MaIAGE, INRAE

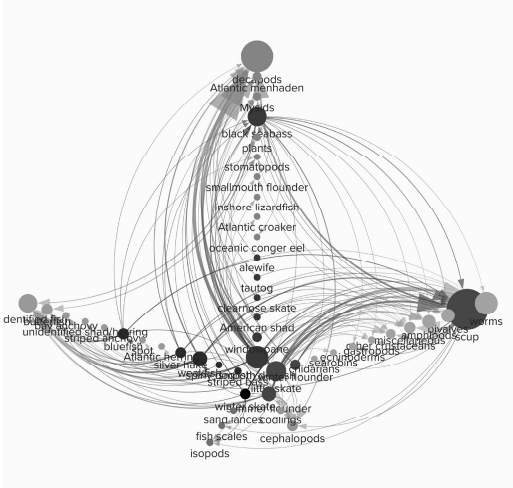
Séminaire d'équipe MSDMA
Conservatoire national des arts et métiers
March 31, 2023

Foodwebs I



Source: www.pw.live/chapter-our-environment/food-web

Foodwebs II



Fall 2021 NEAMAP Food Webs – Mid-Atlantic Coastal Food Web

Source: Virginia Institute of Marine Science

Foodwebs III

Mangal database

- collection of foodwebs on a planetary scale
- 1.300 networks, 120.000 interactions across 7.000 taxa
- contains further information on
  - type (predation, mutualism, parasitism)
  - geographic location
  - climate conditions
- Poisot et al. (2016) + R package `rmangal`

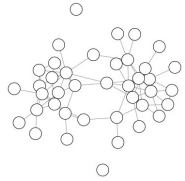
Foodwebs IV

Questions

- analyze and compare many networks
- identify networks with similar structure/organization
- detect outliers/changes...

↪ do it in an **automatic** way
↪ use an **objective criterion** for comparison

Graph clustering task



From a **mathematical point of view**.

- consider a **collection of networks** or graphs
- goal: graph **clustering**
- many other fields of application : social sciences, transport, biology and medicine (e.g. metabolic networks)

Clustering in machine learning literature I

Classical clustering approaches

- clustering of **vectors**: kmeans, DBSCAN, GMM
- straightforward solution for networks (Botella et al., 2022):
  - compute a **graph embedding** (based on hand-designed statistics or representation learning)
  - apply classical ML clustering method