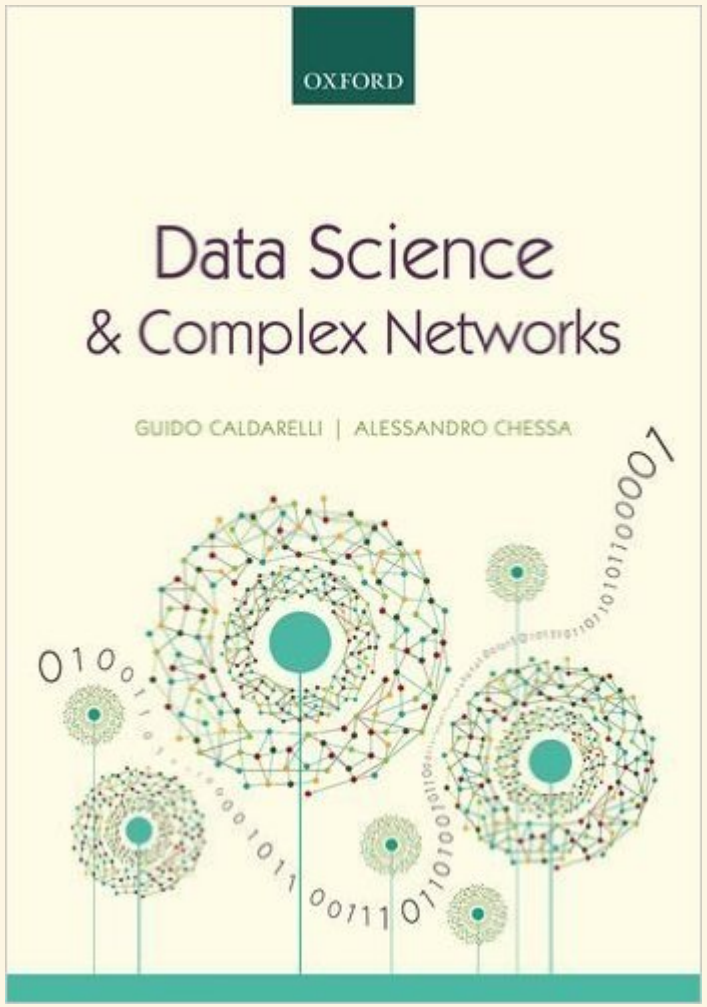
FOOD WEBS

AP

DATA SCIENCE AND COMPLEX NETWORKS



BASIC CONCEPTS

The study of how objects(entities) connect to each other and the properties of their connection.

Possible understanding: a relationship istance.

TERMINOLOGY

- $G = \langle V, E \rangle$ where $E \subseteq V \times V$
- $|V| = n$, $|E| = m$, density is $|E|/|V|^2$
- vertex v is adjacent to u if $(u, v) \in E$; $(v, v) \notin E$.
- neigh. of v, N(v): the set of adjacent vertices; $deg(v) = |N(v)|$
- The adjacency matrix $A_{n \times n}$ of G: $(u, v) \in E \leftrightarrow a_{ij} = 1$
- [incidence matrix $I_{n \times m}$ of G:]

PATHS, CONNECTEDNESS

- A path $u \to v$ is a *sequence* of edges $\langle (u, c_1), (c_1, c_2), \ldots (c_k, v) \rangle$
- its lenght (k+1) is the cardinatility of the path.

- Two vertices are connected if $\exists$ a path betw. them.
- A graph is connected if all its vertices are.

DISTANCES, I

- Distance is the lenght of the (possibly non-unique) shortest path connecting them, ∞ otherwise.
- The diameter of a graph is the maximum distance between any two pairs

DISTANCES, II

Average distances are also important

> […] the first world-scale social-network graph-distance computations, using the entire Facebook network of active users (~721 million users, ~69 billion friendship links). The average distance […] is 4.74, corresponding to 3.74 intermediaries or “degrees of separation.”

WEIGHTED, DIRECTED, MULTIPLEX

$G = \langle V, E, w \rangle$ where $w : E \to \mathbb{R}$

Path lenght: sum of the weights of the arcs.

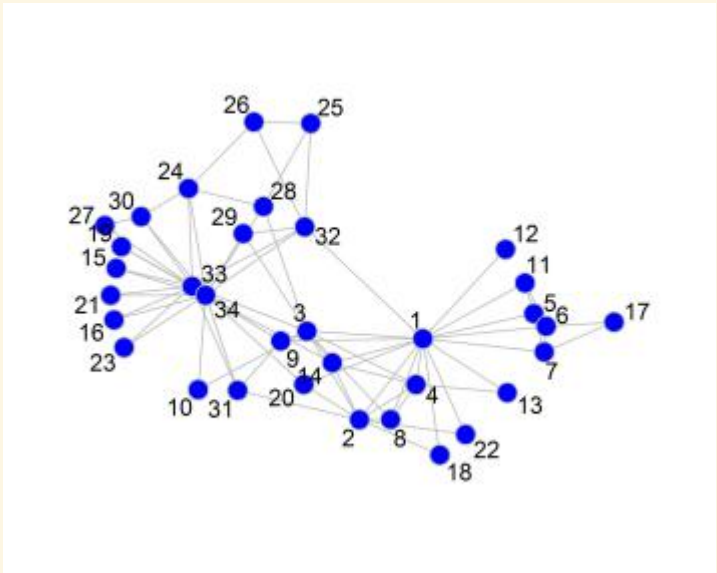
$G = \langle V, E \rangle$ where $v \in V$ are **nodes** and $\langle u, v \rangle \in E$ are **arcs.** $w : E \to \mathbb{R}$

Out-neigh. and In-neigh.

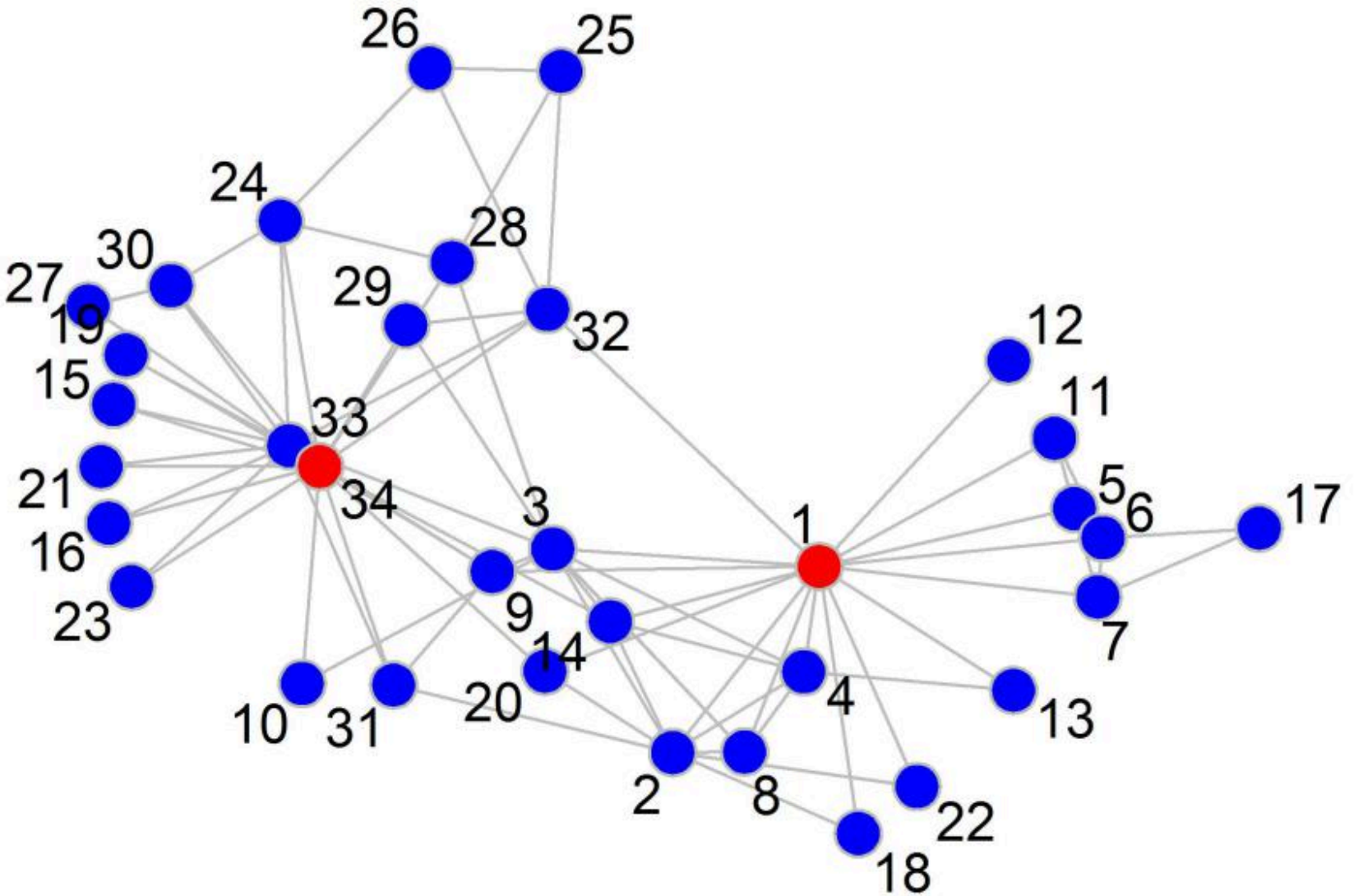
$G = \langle V, E, D \rangle$ where V are **nodes** D are **dimensions/layers** and $\langle u, v, d \rangle \in E$ are **arcs**

INTERESTING QUESTIONS

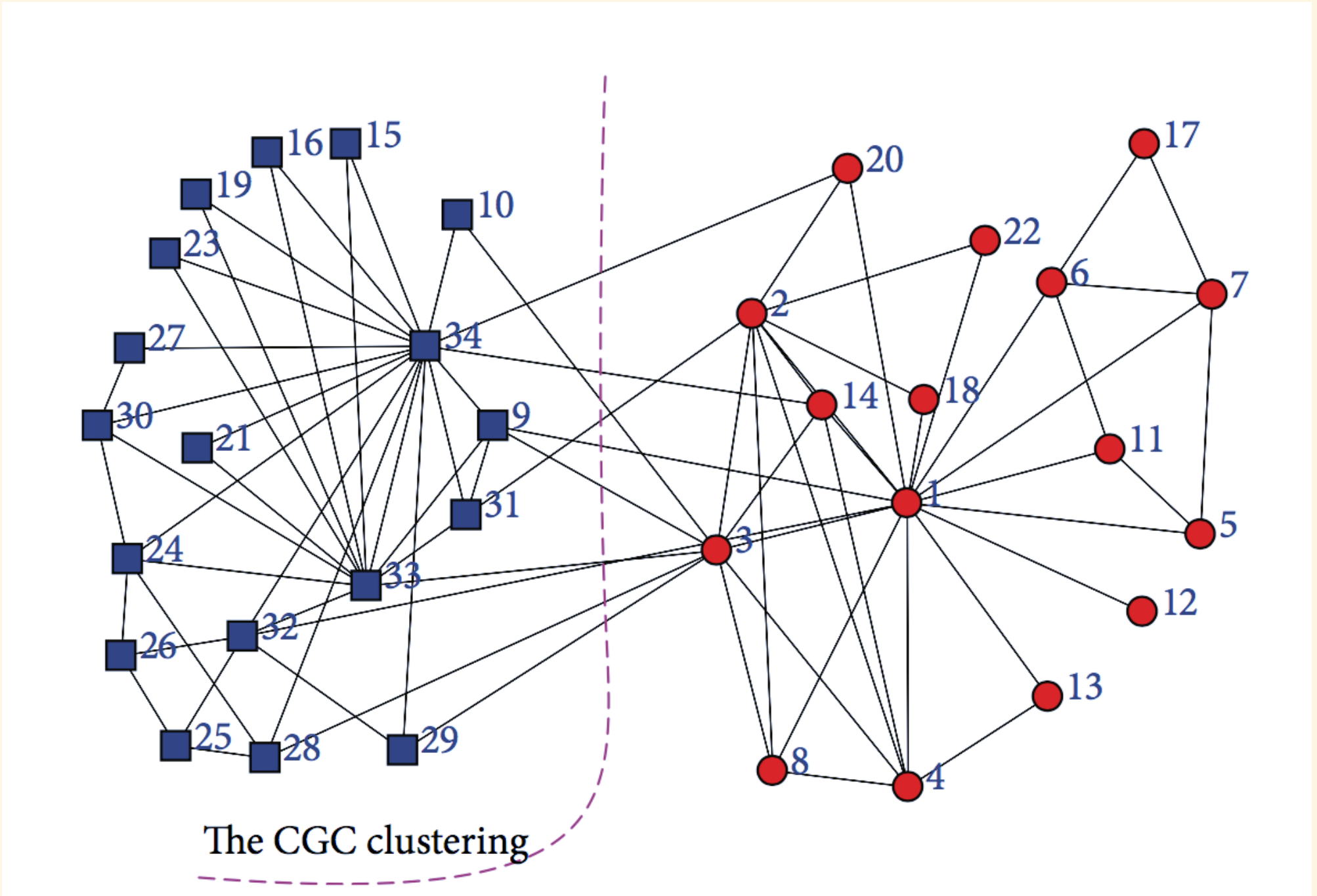
INTERESTING QUESTIONS I: PATHS AND *OFLOW*



INTERESTING QUESTIONS II: CENTRALITIES


26
25
24
28
30
27
29
19
32
12
15
11
33
21
5
34
3
6
17
1
16
9
7
23
14
4
10
31
20
13
2
8
22
18

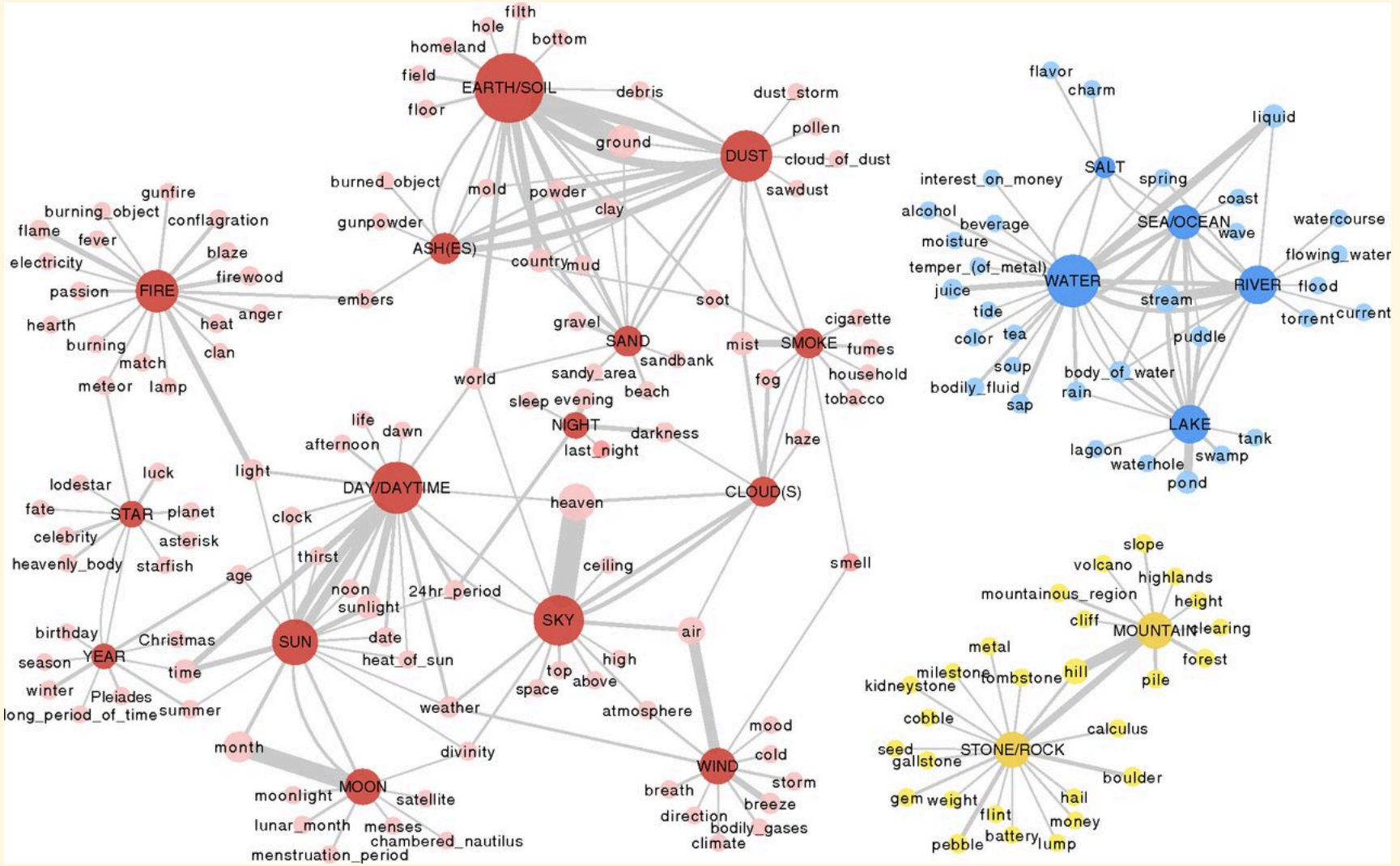
INTERESTING QUESTIONS III: CLUSTERING


16
15
19
10
23
20
17
22
6
7
2
27
34
14
18
30
9
21
11
1
31
5
24
3
33
12
32
26
13
25
28
29
8
4
The CGC clustering

```python
import networkx as nx

G = nx.karate_club_graph()

print("Node Degree")

for v in G:
    print('%s %s' % (v, G.degree(v)))
```

EXTRA: VISUALIZING FREQUENCIES BY GRAPHS

Polisemy:

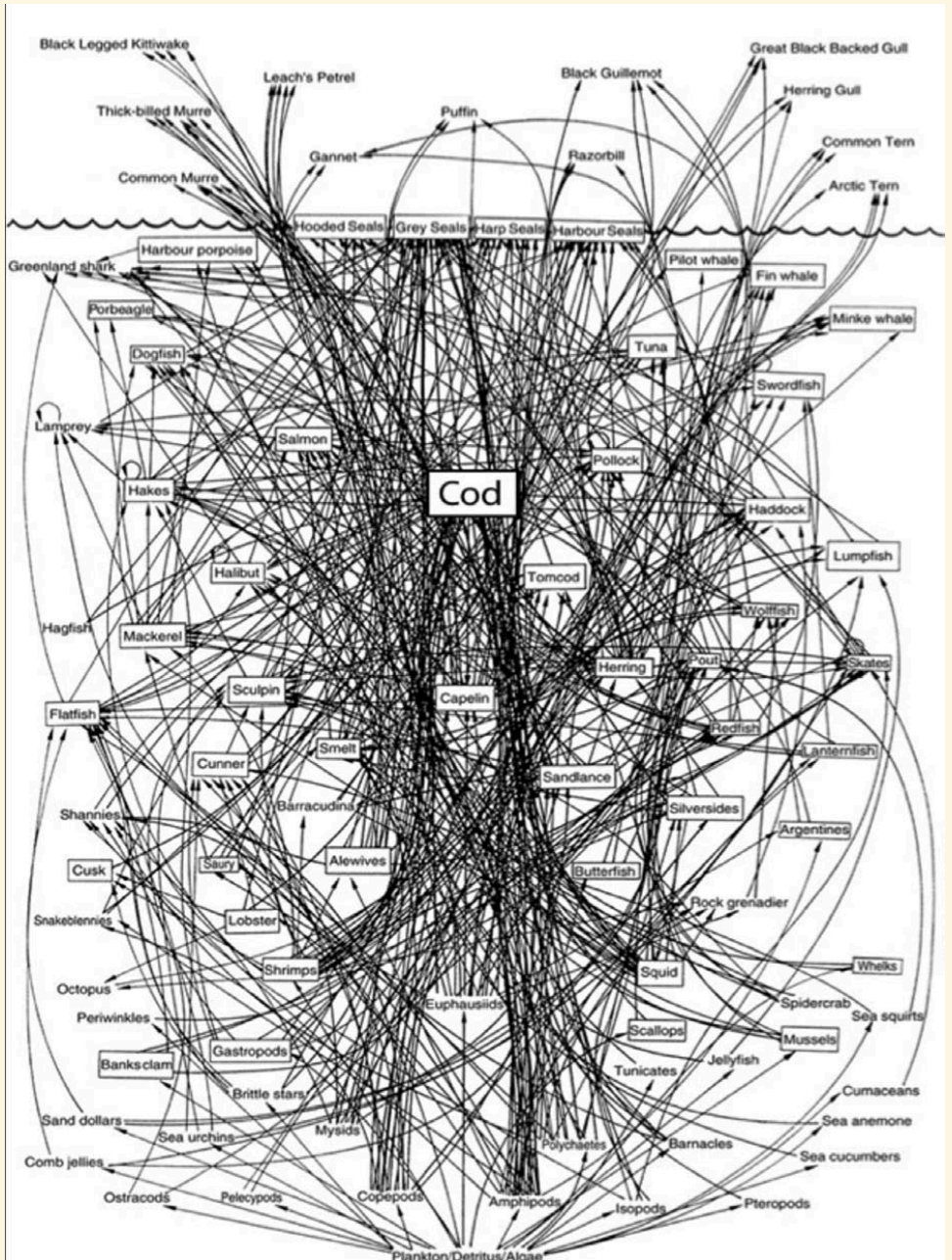


CH. 1: FOOD WEBS

IMPORTANT CONCEPTS

- load data and organise it in a **networkx** data structure
- modeling tip: it is ok to have a special node representing “nature”
- modeling tip: look for *invariants*
- find the connected component (the **bowtie**):

source, connect and sink.



A simplified food web for the Northwest Atlantic. © IMMA

LEARN NETWORK SCIENCE

METHOD

- study *degree distribution*
- find properties of a network in terms of the *degree organization*
- study *clustering coefficient*: why is it so much better than plain network density?

Clone the original (Python 2) from Github

From the same author, a summary of the main concepts