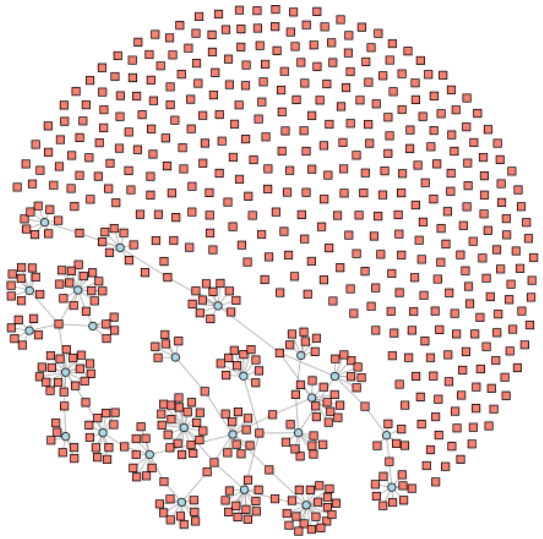
CS4423: Networks

Week 1, Lecture 1: Introduction to CS4423

Dr Niall Madden

School of Mathematical and Statistical Sciences, University of Galway

Wednesday, 15 Jan 2025



This version of the slides are by Niall Madden, but are adapted from original notes by Dr Angela Carevale

Outline

CS4423–Networks is a Semester 2 module on **Network Science**.

Modern societies are in many ways **highly connected**. Certain aspects of this phenomenon are frequently described as **networks**. CS4423 is an introduction to this emerging interdisciplinary subject.

We'll covers several major topics in this module, including

- Graphs and Graph Theory, and how the relate to networks;
- Representations of networks, including as matrices;
- Computing with networks, using `networkx` in Python;
- Centrality Measures
- Random Graphs
- Small Worlds
- Models of Growing Graphs

Lecturer: **Dr Niall Madden** (he/him)

Addressed: Niall (pronounced "Knee"-"al" `#StartsWithAName`)

School: Mathematical and Statistical Sciences, University of Galway.

Office: Room AdB-1013, Arás de Brún

Email: `Niall.Madden@UniversityOfGalway.ie`. This is the best way to contact me. When you do so, please include "`CS4423`" in the subject line. It can also be helpful to include your ID number.

Web: `https://www.niallmadden.ie`

My area of research is **numerical analysis** (mainly differential equations and linear algebra) and **computational mathematics**. Overlap with network science includes use of graph algorithms for matrix partitioning.

There 99 students enrolled in CS4423, from the following programmes:

- Third Arts
- 4th Mathematics and Education
- 4th Computer Science
- 4th Math Science
- **4th Science**
- **4th Financial Mathematics and Economics**
- Structured PhD
- *anyone else?*

The point is: this group comes from very different backgrounds, with different levels of focus on mathematics and computing...

	Mon	Tue	Wed	Thu	Fri
9 – 10			**Lecture**		
10 – 11			**Tutorial/Lab**		
11 – 12					
12 – 1					
1 – 2					
2 – 3				**Lecture**	
3 – 4					
4 – 5		**Tutorial/Lab**			

- Lectures: Wednesday at 9 in McMunn; Thursday at 2 in Larmor;
- Tentative schedule for Labs/tutorials: Tuesday at 4 in AC215 and Wednesday at 10 in CA116a.
- You should be able to attend **one** "lab" session per week; if not, let's discuss.

- There will be some practicals during the semester:
  - Week 3: "Introduction to Python and Jupyter" sessions.
  - Later weeks: help with assignments, prepare for exam, etc.

Quick surveys

1. Who can't attend either Lab Session: **Tue at 4** or **Wed at 10**?
2. Which would you more likely attend: **Tue at 4** or **Wed at 10**?
3. You need your own laptop for labs (but no software). That an issue for anyone? (Chat with me after the class if you prefer).

These notes and supporting Jupyter notebooks, and other relevant material, will be available from the Canvas page:
https://universityofgalway.instructure.com/courses/31889

Usually, notes for the week will be posted on Tuesday evening.

Useful books

- Vito Latora, Vincenzo Nicosia and Giovanni Russo. **Complex Networks.** Cambridge UP 2017.
- Mark Newman. **Networks: an Introduction.** Oxford UP 2018.
- Albert-László Barabási. **Network Science.** Cambridge UP 2016.
- Ernesto Estrada and Philip A. Knight. **A first course in network theory.** Oxford UP 2015.
- David Easley and Jon Kleinberg. **Networks, Crowds and Markets.** Cambridge UP 2010.

- Two homework assignments. Tentative deadlines: Weeks 5 and 10. Will contribute 10% each to the final grade.
- One in-class test. Probably Week 7 (depends on FYP deadlines...). Contributes 10% to the final grade.
- Final exam: 70%.

Examples

Newman (for example) broadly divides the most commonly studied real world networks into four classes:

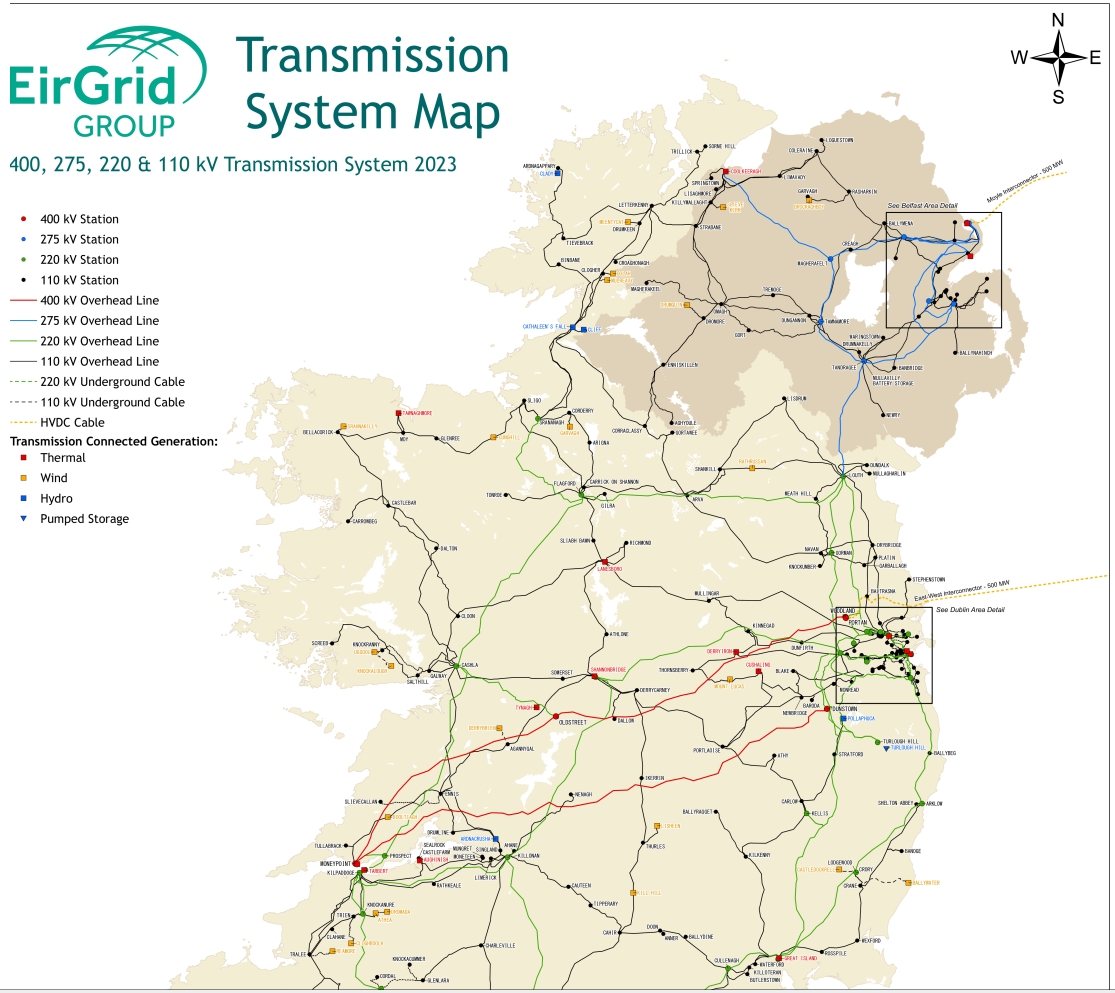
1. **technological** networks,
2. **social** networks,
3. **information** networks and
4. **biological** networks.

- In each case, a network connects parts of a system (**nodes**) by some means (**links**).
- Different techniques used to **display**, **discover** and **measure** the structure in each example.

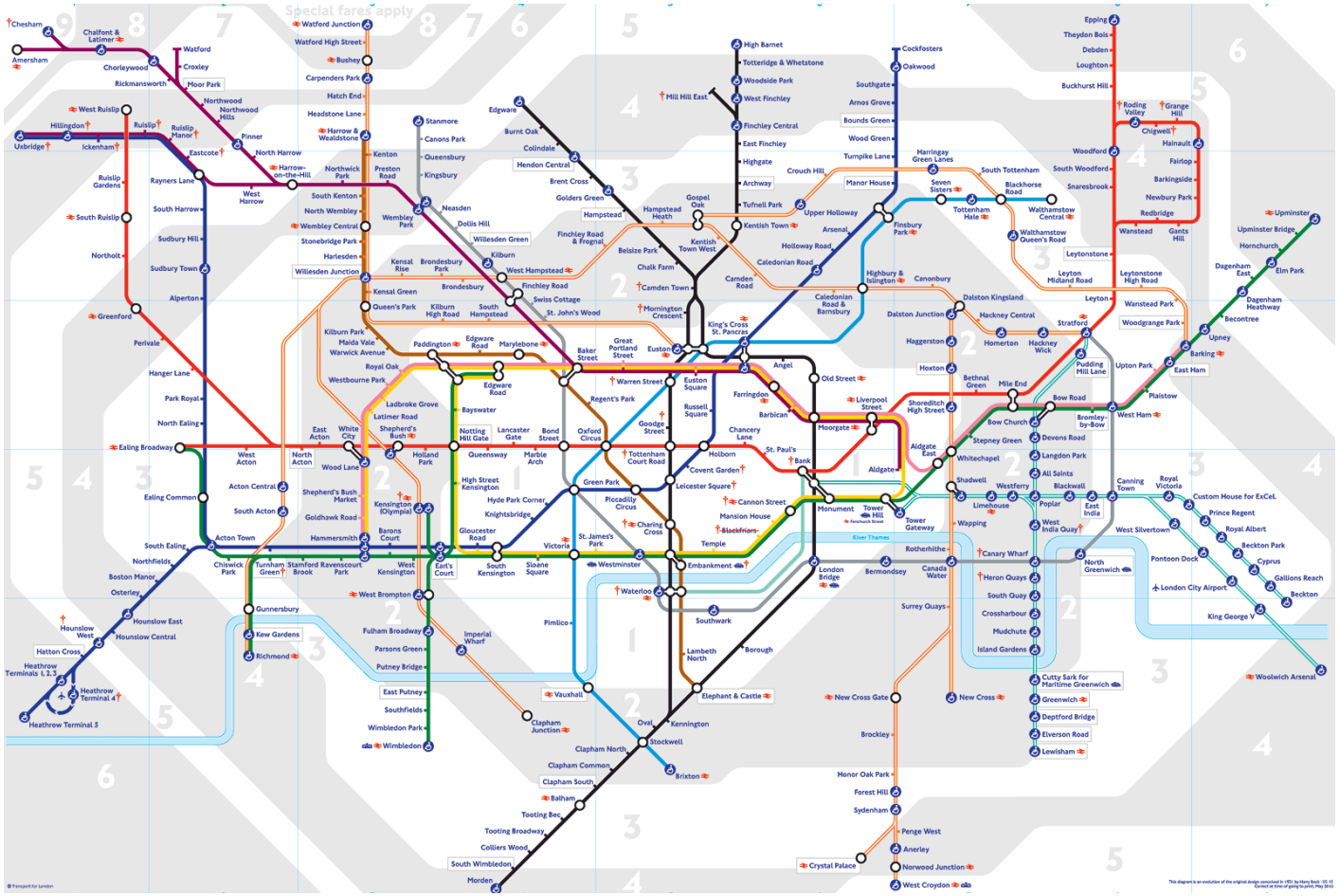
- Technological networks rely on a **physical infrastructure**.
- In many cases, this infrastructure has been built over many decades
- It forms part of the **backbone** of modern societies.
- This includes
  - **road** and other **transportation networks**,
  - **power grids**,
  - **communications networks**

The Irish Power Grid

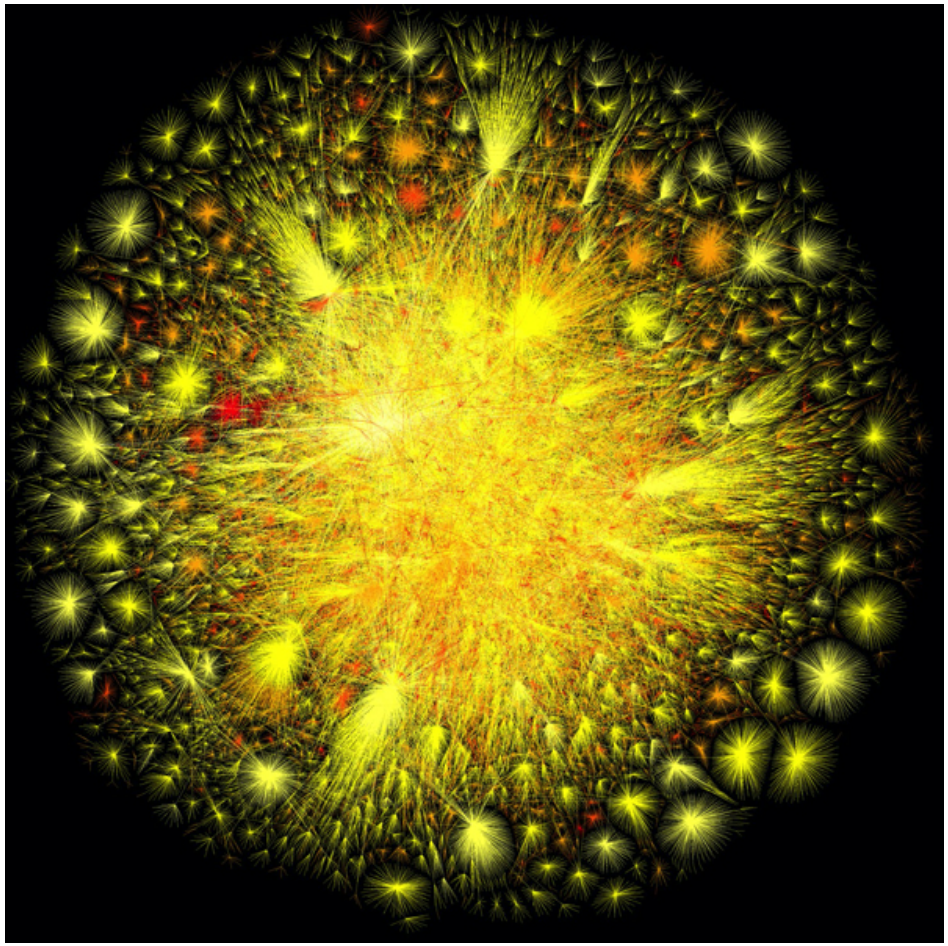
https://www.eirgrid.ie/grid/how-grid-works/irelands-grid



The London Tube



The Internet (circa 2010: opte.org)



- The vertices of a social network are **people** (or, at least, User IDs), with edges representing some sort of **social interaction**.
- In **sociology**, the vertices are often called **actors**, and the edges are called **ties**.
- A social network's **existence** does not depend on social networking sites like `facebook` or `X`.
- Sociologists have studied social networks long before people started exhibiting their relations to others online.
- Traditionally, data about the structure of social networks have been compiled by **interviewing** the people involved.

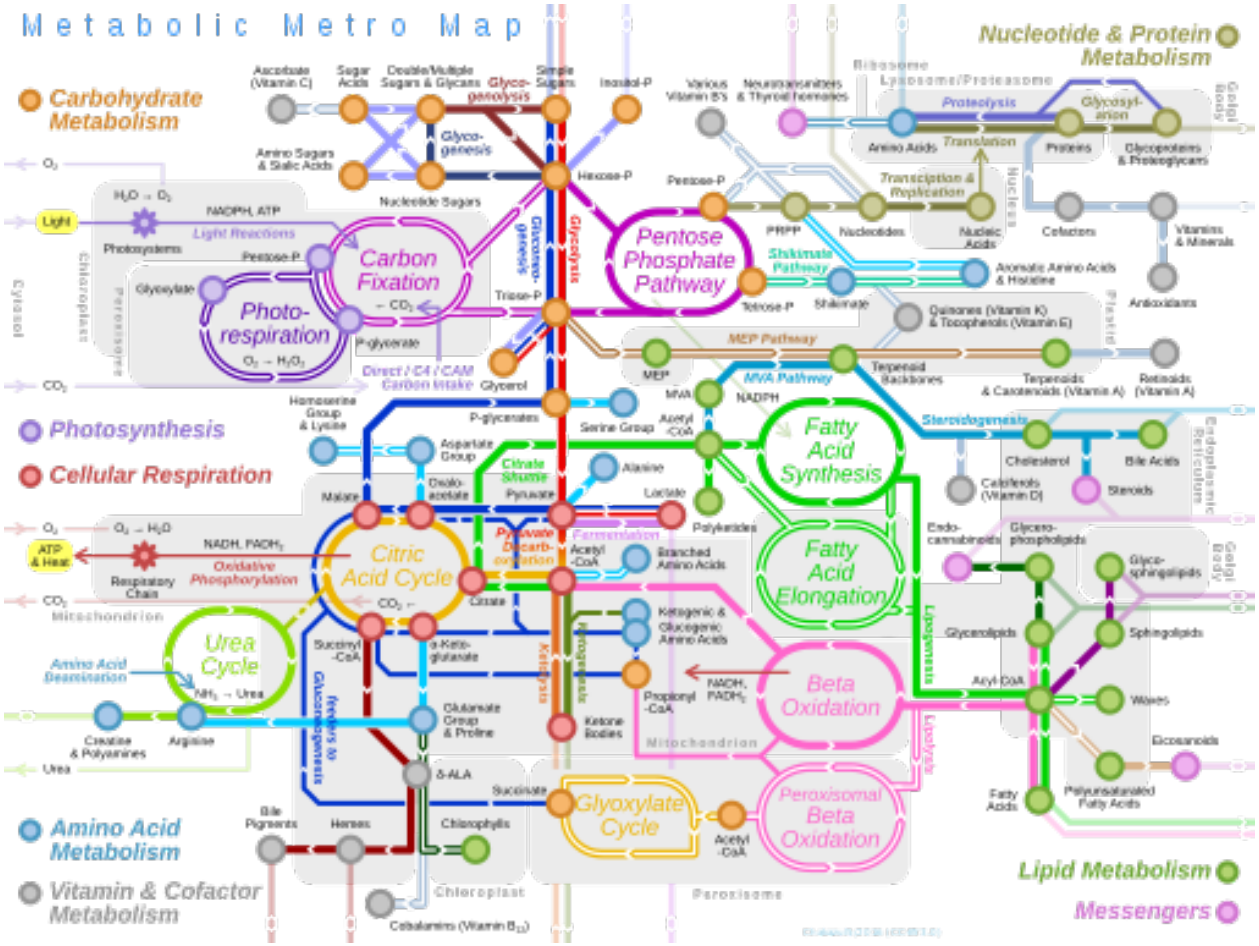
- An **information network** consists of **data items** which are linked to each other in some way.
- Example: **relational databases**.
- Sets of information (like **scientific publications**) have been linking to each other (e.g., through **citations**) long before computers were invented.
- However, links in digital form are easier to follow ...

The (World Wide) Web

- The **WWW** is probably the most wide spread and best known examples of an information network.
- Its nodes are **web pages** containing information in form of text and pictures, and its edges are the **hyperlinks**, allowing us to surf, or **navigate** from page to page.
- Hyperlinks run in **one direction only**, from the page that contains the hyperlink to the page that is referenced.
- Therefore, the WWW is a **directed network**, a graph where each edge has a direction.

- **Biochemical networks** represent molecular level patterns of interaction and control mechanisms in the biological cell:
  - **metabolic** networks,
  - **protein-protein interaction** networks and
  - **genetic regulatory** networks.
- A **neural network** can be represented as a set of vertices, the neurons, connected by two types of directed edges, one for excitatory inputs and one for inhibiting inputs. (Not to be confused with an **artificial neural network**).
- **Ecological networks** are networks of ecological interactions between species.

Metabolic Networks



(source: Wikipedia)

In its simplest form, a network is just a collection of points (called **vertices** or **nodes**), some of which are joined in pairs (called **edges** or **links**).

- Many systems of interest are composed of individual parts that are in some way linked together.
- Such systems can be regarded as networks, and thinking about them in this way can often lead to new and useful **insights**.
- Network science studies the **pattern of connections** between the components of a system.
- Naturally, the **structure** of the network can have a big impact on the **behavior** of a system.

- A network is a **simplified representation** of a complex system by vertices and edges.
- The scientific study of networks is an **interdisciplinary** undertaking that combines ideas from
  - mathematics,
  - computer science,
  - physics,
  - the social sciences,
  - biology
  - ...
- Between these scientific fields, many **tools** have been developed for **analyzing**, **modeling** and **understanding** networks.

Centrality is an example of a useful and important type of network measures. It is concerned with the question of how **important** a particular vertex or edge is in a networked system.

Different concepts have been proposed to capture **mathematically** what it means to be central.

Example: a simple measure of the centrality of a vertex is its **degree**, i.e., the number of edges it is part of (or equivalently the number of vertices it is adjacent to).

Applications: determine which entities in a social network have the most influence, or which links in a power grid are most vulnerable.

Which measurements and calculations give **meaningful answers** for a particular system depends of course on the specific nature of the system and the questions one wants to ask.

- Another interesting network concept is the **small-world effect**.
- It is concerned with the question of how far apart two randomly chosen points in a network typically are.
- Here, **distance** is usually measured by the number of edges one would need to cross over, when travelling along a **path** from one vertex to the other.
- In real world social networks the distance between people tends to be rather small.
- This observation is known as the **six degrees of separation** in popular culture.

Jupyter Notebooks and `networkx`

- Lecture notes and Assignments come in the form of `jupyter` notebooks.
- This allows us to include interactive `python` code with the text.
- The notebooks will be uploaded on the CS4423 web page, together with a PDF printout.
- You can access Jupyter at `https://cloudjupyter.universityofgalway.ie/` (more info tomorrow). Or you can install it on your own computer.
- Alternatively, you can use services such as **binder** or Google's Colab.