COMP2200/COMP6200 – Week 11B Paths and centrality

Greg Baker

21st October 2025

Linking concept to intuition

Where we are headed

- Question from last slide: why do we care who sits at the centre of the web?
- Today's mission: build intuition for shortest paths, then translate that into Orange widgets and a small Python check.
- Keep the study group network open. We will keep poking it from different angles.

Shortest path in plain words

- Imagine messaging across a friendship chain. How many forwards until it arrives? That count is the **path length**.
- Dijkstra's algorithm repeats a very human process: visit the closest unvisited friend, update what you know, repeat.
- Modern systems often use heuristics or parallel search, but this classic still builds solid intuition.

Try it manually

Pick two students from the data set. Trace the minimal steps on paper before Orange does it for you.

Orange workflow

Orange demo: Network Analysis widget

1. Add **Network Analysis**. Choose metrics: degree, betweenness, closeness, eigenvector.
2. Push results into a **Data Table**. Notice the new columns prefixed with `NA-`.
3. Sort by each metric. Who ranks highest and why? Does it match your intuition from the sketch?

Reading the metrics

- **Degree centrality**: How many direct study buddies do you have?
- **Betweenness**: How often are you on the critical bridge between others?
- **Closeness**: On average, how many study hops to reach everyone else?
- **Eigenvector/PageRank flavour**: Are your friends themselves influential?

Reflection prompt

Which metric best captures "most helpful peer mentor" in our context?

Python side quest

Checking Orange with NetworkX

```
import pandas as pd
import networkx as nx

g_edges =
    pd.read_csv("resources/study_group_edges.csv")
G = nx.from_pandas_edgelist(g_edges,
    "source", "target",
    edge_attr="study_hours")

betweenness = nx.betweenness_centrality(G,
    weight="study_hours")
for name, score in betweenness.items():
    print(name, f"{score:.3f}")
```

- Run this to sanity-check Orange's betweenness numbers.
- Expect tiny differences due to normalisation choices

Mini challenge: identify a bridge

- Find the student whose betweenness is highest. Why do they matter for keeping the group connected?
- If that student drops the unit, what happens to average path length? Discuss with your neighbour.
- Post your hypothesis on Ed after class. I'll respond with a quick NetworkX notebook showing the impact.

Take a breather

- We now have a toolkit for path length intuition plus degree/bridging centrality.
- Next up: zooming into ego networks and spotting communities.
- Any questions before we stretch our legs?