The Navigation Problem

- You are an individual (vertex) in a very large social network
- You want to find a (short) chain of friendships to another individual
- You don’t have huge computers and a global/bird’s-eye view
- All you (hopefully) know is who your neighbors/friends are
  - ...and perhaps information about them (age, interests, religion, address, job,…)
- You can ask your friends to make introductions, which lead to more
- How would you do it?
- Also known as search in networks and the “small world problem”
- Small diameter is necessary but not sufficient!
  - ...navigation is an algorithmic problem
- Related to the problem of routing data packets in the Internet
Small Worlds and the Law of the Few

- Travers & Milgram 1969: classic early social network study
  - destination: a Boston stockbroker; lived in Sharon, MA
  - sources: Nebraska stockowners; Nebraska and Boston “randoms”
  - forward letter to a first-name acquaintance “closer” to target
  - target information provided:
    - name, address, occupation, firm, college, wife’s name and hometown
    - navigational value?
- Basic findings:
  - 64 of 296 chains reached the target
  - average length of *completed* chains: 5.2
    - interaction of chain length and navigational difficulties
  - main approach routes: home (6.1) and work (4.6)
  - Boston sources (4.4) faster than Nebraska (5.5)
  - no advantage for Nebraska stockowners
The Connectors to the Target

- **T & M** found that many of the completed chains passed through a very small number of penultimate individuals
  - Mr. G, Sharon merchant: 16/64 chains
  - Mr. D and Mr. P: 10 and 5 chains

- Connectors are individuals with extremely high degree
  - why should connectors exist?
  - how common are they?
  - how do they get that way? (see Gladwell for anecdotes)

- Connectors can be viewed as the “hubs” of social traffic

- Note: no reason *target* must be a connector for small worlds

- Two ways of getting small worlds (low diameter):
  - truly random connection pattern → dense network
  - a small number of well-placed connectors in a sparse network
Small Worlds: A Modern Experiment

• The Columbia Small Worlds Project:
  – considerably larger subject pool, uses email
  – subject of Dodds et al. assigned paper

• Basic methodology:
  – 18 targets from 13 countries
  – on-line registration of initial participants, all tracking electronic
  – 99K registered, 24K initiated chains, 384 reached targets

• Some findings:
  – < 5% of messages through any penultimate individual
  – large “friend degree” rarely (< 10%) cited
  – Dodds et al: \(\rightarrow\) no evidence of connectors!
    • (but could be that connectors are not cited for this reason…)
  – interesting analysis of reasons for forwarding
  – interesting analysis of navigation method vs. chain length
The Strength of Weak Ties

• Not all links are of equal importance
• Granovetter 1974: study of job searches
  – 56% found current job via a personal connection
  – of these, 16.7% saw their contact “often”
  – the rest saw their contact “occasionally” or “rarely”
• Your “closest” contacts might not be the most useful
  – similar backgrounds and experience
  – they may not know much more than you do
  – connectors derive power from a large fraction of weak ties
• Further evidence in Dodds et al. paper
• T&M, Granovetter, Gladwell: multiple “spaces” & “distances”
  – geographic, professional, social, recreational, political,…
  – we can reason about general principles without precise measurement
The Magic Number 150

- Social channel capacity
  - correlation between neocortex size and group size
  - Dunbar’s equation: neocortex ratio → group size
- Clear implications for many kinds of social networks
- Again, a *topological* constraint on typical degree
- From primates to military units to Gore-Tex
Summary, and a Mathematical Digression

• So far:
  – large-scale social networks reliably have high-degree vertices
  – large-scale social networks have small diameter
  – furthermore, people can find or navigate the short paths from only local, distributed knowledge
  – these properties are true of other types of networks, too

• But there must be some limits to degrees
  – can’t be “close friends” with too many people (150? 1000?)

• Large N, small diameter and limited degrees are in tension
  – not all combinations are possible

• Let N be population size, Delta be the maximum degree, and D be the diameter
• If Delta = 2 then must have \( D \sim N/4 \) (\( >> 6, >> \log(N) \))
Summary, and a Mathematical Digression

• The relationship between D, Delta and N has been studied mathematically
• For fixed D and Delta, largest N can be is

\[ N \leq \Delta^D \]

• For example: if N = 300M (U.S. population) and Delta = 150, get constraint on D:

\[ 300,000,000 \leq (150)^D \]
\[ \log(300,000,000) \leq D \log(150) \]
\[ D \geq 3.9 \]

• So calculation consistent with reality (whew!)
• More generally: multiple structural properties may be competing
Two Aspects of Navigation

• In order for people (or machines) to find short paths in networks:
  – short paths must exist (structural; small diameter)
  – people must be able to find the short paths via only local forwarding (algorithmic)

• The algorithmic constraints are strong (Travers & Milgram)
  – only know your neighbors in the network
  – limited information about the target/destination (physical location, some background)

• Look at a model incorporating structural and algorithmic constraints
Kleinberg’s Model

• Start with an $k$ by $k$ grid of vertices (so $N = k^2$)
  - each vertex connected to compass neighbors
  - add a few random “long-distance” connections to each vertex
  - probability $p(d)$ of connecting to a vertex at grid distance $d$:
    \[ p(d) \propto \left(\frac{1}{d}\right)^r, r \geq 0 \]
  - large $r$: heavy bias towards “more local” long-distance connections
  - small $r$: approach uniformly random
Kleinberg’s Question

• Which values of $r$:

$$p(d) \propto (1/d)^r, r \geq 0$$

permit efficient navigation?

• Efficient: number of hops $<< N$, e.g. $\log(N)$

• Algorithmic assumption:
  – vertices know the grid addresses of their neighbors
  – vertices know the grid address of the target (Sharon, MA)
  – vertices always forward the message to neighbor closest to the target in grid distance
  – no “backwards” steps, even if helpful
  – purely geographic information
Kleinberg’s Result

- **Intuition:**
  - if $r$ is too *large* (strong local bias), then “long-distance” connections never help much; short paths may not even *exist*
  - if $r$ is too *small* (no local bias), we may quickly get close to the target; but then we’ll have to use grid links to finish
  - effective search requires a delicate *mixture* of link distances

- **The result (informally):** as $N$ becomes large:
  - $r = 2$ is the *only value* that permits rapid navigation ($\sim \log(N)$ steps)
  - a “knife’s edge” result; very sensitive

- **Note:** *locality of information* crucial to this argument
  - At $r \leq 2$, will have small diameter, but local algorithms can’t find the short paths
<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Distance</th>
<th>Time</th>
<th>Average Speed</th>
<th>Social Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-Mar-2023 08:54</td>
<td>Jacksonville, FL</td>
<td>12.2 miles</td>
<td>27 min, 37 sec</td>
<td>50 mph</td>
<td>6.3</td>
</tr>
<tr>
<td>23-Mar-2023 06:19</td>
<td>Jacksonville, FL</td>
<td>11.1 miles</td>
<td>20 min, 33 sec</td>
<td>50 mph</td>
<td>6.9</td>
</tr>
<tr>
<td>21-Mar-2023 05:05</td>
<td>Jacksonville, FL</td>
<td>10.2 miles</td>
<td>18 min, 49 sec</td>
<td>50 mph</td>
<td>6.4</td>
</tr>
</tbody>
</table>

The bill has traveled at least 6,000 miles since it was introduced in Congress. It continues to move at an average speed of 50 miles per day.
From Brockmann, Hufnagel, Geisel (2006)  
Best-fit value of $r = 1.59$
Navigation via Identity

- Watts et al.:
  - we don’t navigate social networks by purely “geographic” information
  - we don’t use any *single* criterion; recall Dodds et al. on Columbia SW
  - different criteria used at different points in the chain

- Represent individuals by a *vector* of attributes
  - profession, religion, hobbies, education, background, etc…
  - attribute values have distances between them (tree-structured)
  - distance between individuals: minimum distance in *any* attribute
  - only need *one thing in common* to be close!

- Algorithm:
  - given attribute vector of target
  - forward message to neighbor closest to target

- Let’s look a bit at the [paper](#)

- Permits fast navigation under broad conditions
  - not as sensitive as Kleinberg’s model
Summary

- Efficient navigation has both structural and algorithmic requirements
- Kleinberg’s model and question captures both
- Result predicts delicate mixture of connectivity for success
- Not too far from reality? (Where’s George? data)
- Watts et al. provide more “sociological” model
- More complex, but less sensitive