Lecture 1: Connect (1/4)

How the friendship we form connect us? Why we are within a few clicks on Facebook?

COMS 4995-1: Introduction to Social Networks
Tuesday September 4th
You have to sign up yourself in piazza!
  o If you don’t want to use your columbia account that’s ok, but you need to ask me to sign you up.
  o BTW: no need to sign up for the wiki, it’s open

You have to complete two enrolling steps:
  o Sign the cover sheet (acknowledging our rules)
  o Complete our preliminary survey (contact us if you want to do it anonymously)

You should plan warming up for the 1st assignment
What is the small world effect?

“You’re in NY, by the way, do you happen to know x?
- In fact, yes, what a small world!”

What seems remotely distant is indeed socially close.

The small world problem, S. Milgram, Psychology today (1967)
Why study small world?

- Small world hypothesis is *necessary* for
  - Fast and wide information propagation
  - Tipping point: small change has large effect
  - Network robustness and consistency
- Studying the conditions of the “small world” effects tells us a lot about *how* we are connected?
- “Man is, by nature, an animal of a society”
  “ζωὸν πολιτικὸν”
Small world: a simplistic argument

* Remember how the King lost his fortune to the chess player?
* What would you turn down an offer?
  - An daily doubling series \{1¢, 2¢, 4¢, etc.\} over a month
  - Against $1,000? YES/NO
  - Against $100,000? YES/NO
  - Against $1,000,000? YES/NO
  - Against $100,000,000? YES/NO
### Small world: a simplistic argument

- **How many people would you recognize by name?**
  - ‘67 M. Gurevitch (MIT): about 500

- **Roughly, how many are socially related to you?**

<table>
<thead>
<tr>
<th>how close to you?</th>
<th>Compares to</th>
<th>%. US pop.</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 direct acquaintance</td>
<td>C.S. dept</td>
<td>0.00017%</td>
</tr>
<tr>
<td>250,000 share an acquaintance with you</td>
<td>Harlem district</td>
<td>0.083%</td>
</tr>
<tr>
<td>125m share an acquaintance with a friend of yours</td>
<td>Northeast + Midwest</td>
<td>42%</td>
</tr>
</tbody>
</table>
Small world: The skeptics

* The previous model is way too optimistic!

- Reason #1: it assumes acquaintance set are disjoint
  - whereas they are related as part of a social graph
  - expansion through the graph may be limited.
The previous model is way too optimistic!

- Reason #2: social acquaintances are biased
  Geography, occupation & social status, race
  * Favors clusters, inbreeding. Increases social distance
- May be others: a few know much more people
Outline

* Milgram’s “small world” experiment

* It’s a “combinatorial small world” (Sep. 8th)
* It’s a “complex small world” (Sep. 13th)
* It’s an “algorithmic small world” (Sep. 15th)
The 10 papers that will make you a social expert
A direct experimental approach
- Main issue: we do not know the graph (This is 1967!)
- Can we still find a method to exhibit small chains of acquaintance between two arbitrary people?

At least we can try it out:
- Pick a single “target” and an arbitrary sample of people
  Send a “folder” with target description, asking participants to
- Add her name, and forward to her friend or acquaintance who she thought would be likely to know the target,
- Send a “tracer” card directly to a collection point
Would this work?

First folders arrived using 4 days and 2 intermediaries

2 successful experiments

- 1\textsuperscript{st}: 44 folders out of 160
- 2\textsuperscript{nd}: 64 folders on 296
- Average distance = 6.2.

The small world problem, S. Milgram, Psychology today (1967)
More observations

- Chains not random
  locally: gender, family/friends
  globally: hometown/job

- Some received multiple folders (up to 16)

- Drop out-rate ~25%
  May artificially reduce distance and can be corrected

Are these results reproducible?

- Yes, using new forms of communication

Another method: assuming graph is known

- Exhaustive search:
  - Erdös number, Kevin Bacon number
  - More evidence of short paths, including in different domains
Confirmation using email

- Larger scale
  18 targets, 24,000 starting points

- Drop-out rates: 65%
  success rate is lower: 1.5%
  median distance:
  4 raw data and 7 corrected data

- Role of social status
  success rate changes, but not mean distance

Even today: Sign-up at smallworld.sandbox.yahoo.com

An experimental study of search in global social networks.
Erdös number

Connections = collaboration

- 2 persons connected if they co-authored a paper.

How far are you from Erdös?

- 94% within distance 6
- Median: 5
- Small number prestigious

Field medalists: 2-5; Abel prize: 2-4

The Erdös Number Project, www.oakland.edu/enp/
Kevin Bacon number

Connections = collaboration
- 2 actors connected if they appear in one movie
- ~88% actors connected
- 98.3% actors: 4 and less

Maximum number is 8.

The oracle of Bacon, oracleofbacon.org
Instant Messaging

Different scales

- 180m nodes, 13b edges

Connections = collaboration
- 180m nodes, 1.3b edges

Planetary-scale views on a large instant-messaging network.
J. Leskovec and E. Horovitz. WWW (2008)
Outline

* Milgram’s “small world” experiment
* It’s a “combinatorial small world” (UPCOMING)
* It’s a “complex small world”
* It’s an “algorithmic small world”