Lab #8: Graph Detective

Everything Data

CompSci 216 Spring 2015

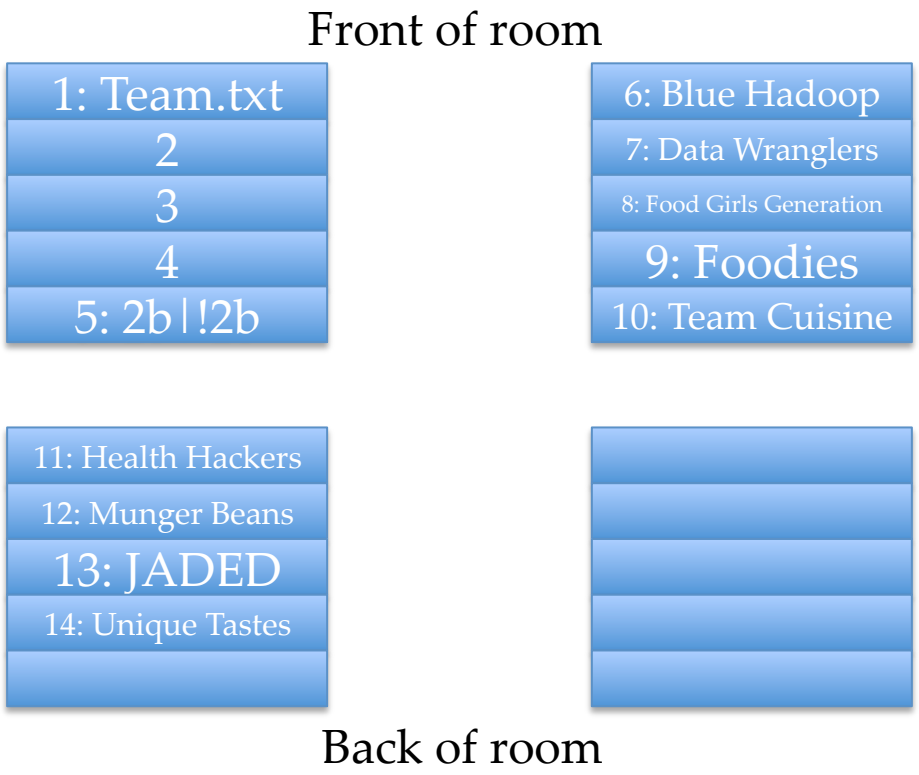
DUKE
COMPUTER SCIENCE

Announcements (Mon. Mar. 16)

- **Homework #8** deadline extended to tonight
  - Because of the extension, Jun will have office hours today after class in LINK until 5:45pm (instead of Friday) this week
- **Project mid-term report** due in two weeks
  - Feedback on your proposal has been already sent during break, or will be by tonight

Seat assignment

See course website (under "Schedule") for project teams



A quick review

Useful stats when getting to know a graph:

- # of nodes and # of edges
- Degree distribution
- Distance distribution
- Clustering coefficients

Some more review

Interesting properties of real graphs

- Power law, small world, triadic closure

Models for generating graphs

- Random, *Watts-Strogatz*, *Barabasi-Albert*
- Keep in mind their properties
  - Feel free to consult Lecture 8 slides

Challenge

- 7 files (.xml.gz)
- 7 graphs
  - Random-Sparse
  - Random-Dense
  - Barabasi-Albert with $m = 1$
  - Barabasi-Albert with $m = 3$
  - Congress
  - Coauthor
  - Facebook
- Which one is which?

http://informoverload.com/wp-content/uploads/2013/11/sep_A.jpg

Rules: getting checked

- Give a "complete" answer
  - Okay to guess those you are not sure about
- You will be told how many correct matches you have, but not which ones!
- Three "lifelines"
  - Each allows you to get a yes/no answer to a question "is our answer for file *i* correct?"

Rules: winning

Highest number of correct matches wins

- If there is a tie, break by # of lifelines used
- If there is still a tie, break by time

Extra credit: 10% for matching all but two

Lessons learned

- Often it's hard to wrap your head around a big graph
- Remember those simple distributions and properties to look for
  - They can be surprisingly effective in painting a "big picture" for you

http://honolulunotes.com/2012/10/08/checking-the-numbers/messy-graph-size/

Finally

- Remember to submit `hw08` and `team.txt` **under `lab08` by midnight**
- Sample solutions to both will be posted by tomorrow night