Internet-scale Experimentation

To do ...

- Welcome
- Class organization and guidance
- Some good advice
The state of affairs

- An ever growing Internet
  - ~3 billion people
  - 15 billion devices connected
  - 10 thousands ISPs
  - >52 thousands networks (ASes)

- Tons of money at play
  - Alphabet (Google’s parent company) 2nd Q 2016 revenues - $21.5 billions (+21% per year)

*Yuhai Tu, How robust is the Internet?, Nature (27 July 2000)
The state of affairs

- Society’s increased dependency on
  - More, ever-larger Internet-scale systems
    - FB, Skype, Twitter, Google, Akamai, Amazon, Netflix …
  - Facebook’s 1.5 billion monthly users
    - Average time in FB 20’/day
    - Or 20% of all online time

- We still
  - Don’t know how the network underneath looks like
  - Can’t predict these systems’ behaviors
  - or trust their security, performance, resilience, …
  - …
Experimentation and measurements

- For answers, validation, new questions - Experiments

- Other reasons for measurement and experimentation
  - Technical – Designing network components and protocols
  - Commercial – Demographic, performance properties
  - Social – Statistics on network activities, characterizing popularity and content

- But not without challenges
  - How to do experimentation at Internet-scale?
  - What’s representative?
  - What’s reproducible?
  - …
Our goal

- To explore the challenges with Internet-scale experimentation
- To learn the strengths/weaknesses of available and proposed tools, platforms and approaches

A higher level goal – learn research by doing
- Reading research papers
- Discussing them in depth
- Presenting research ideas
- Picking a research project and executing it
- Writing a research paper
Class structure

• Meetings
  – Read research papers
  – Present research ideas
  – Learn about ongoing work

• Take part of a mini-conference
  – You as the Program Committee
  – (most of) Our reading list as the submissions
  – Task – review and discuss them in a PC meeting

• Project
  – Come up with fresh idea, we can brainstorm together
  – Research, build, test, …
  – Write a report/paper on your project
  – Present it to the class
What it means to you

- Reading and class discussion (10%)
  - About 34 papers over the quarter (~4 papers per week), but you can “quickly” read most of them

- Writing reviews and discussing papers as part of a program committee (20%)
  - 7 reviews over the quarter

- Presenting papers (20%)
  - 2 paper presentations on each “side” of the midterm

- Working on a project (50%)
What it means to you

- **Reading and class discussion (10%)**
  - About 34 papers over the quarter (~4 papers per week), but you can “quickly” read most of them

- **Writing reviews and discussing papers as part of a program committee (20%)**
  - 7 reviews over the quarter

- **Presenting papers (20%)**
  - 2 paper presentations on each “side” of the midterm

- **Working on a project (50%)**
Why reading?
- For a conference or a class
- To learn a new field
- To keep up to date

Way too many papers
- SIGCOMM 2015: 42, IMC 2015: 44, CoNEXT 2015 ~41
- NSDI, Mobisys, HotNets, PAM, OSDI, ATC …
- Per year >200 papers on your area

You are rarely taught how to do it
A three-pass approach (Keshav’s)

1. Get the general idea
   – 5-10’
   – Good for a paper that’s not in your area

2. Grasp the content
   – ~1hr for an experienced reader
   – Good for a paper that’s interesting but not in your speciality

3. Understand in depth
   – n-hours for beginners; >1-2hr for experienced readers
   – To fully understand a paper; e.g., if you are a reviewer
First pass ...

- Process
  1. Carefully read title, abstract, intro
  2. Read section, subsection headings, ignore the rest
  3. Glance at the math
  4. Read conclusions
  5. Glance at references, mentally ticking off what you have read

- Keep in mind
  - Invariant comprehension questions
    • What is the problem? … contributions? … conclusions? What is the support for the conclusions?
  - Invariant evaluation questions
    • What’s the intellectual nugget? What’s the main conclusion/contribution? Why is that important? Does the content support it?
First pass

- At the end, you should answer
  1. Category – Measurement, analysis of a system, …
  2. Context – Other papers, theoretical foundations
  3. Correctness – Are the assumptions valid?
  4. Contributions – What are the main contributions?
  5. Clarity – Is it well written?

- Do you need to go beyond?

- *BTW, reviewers/readers will do the same with your paper*
Second pass

• Process
  – Read with greater care, avoid details (e.g., proofs)
  – Take down points or make comments in margins
  – Look at figures, diagrams, illustrations
  – Mark relevant unread references for further reading

• At the end, you should be able to
  – Grasp the content
  – Summarize the main thrust with evidence to someone else
Third pass

- (Virtually) re-implement the paper
  - Making same assumption, re-create the work and compare
  - To see the innovations, hidden failings and assumptions
- Identify and challenge every assumption

- Other takeaways
  - Consider how you would present every idea
  - Write down presentations/writing/methodological ideas
  - Potential future work

- At the end you should be able to
  - Reconstruct the entire paper structure from memory
  - Identify the strong/weak points
What it means to you

- Reading and class discussion (10%)
  - About 36 papers over the quarter (~4 papers per week), but you can “quickly” read most of them

- Writing reviews and discussing papers as part of a program committee (20%)
  - 7 reviews over the quarter

- Presenting papers (20%)
  - ~2 paper presentations on each “side” of the midterm

- Working on a project (50%)
Taking part of a mini-conference as a PC member
  - Practice reviewing papers

How does it work?
  - State your preferences (in HotCRP … soon)
  - Papers assigned based on them
  - Each paper will receive 3 reviews, you will write 7 (28 papers * 3 reviews / 11 students + me)
  - Discuss papers in a two-part PC meeting

To do next – read the linked papers on reviewing
  - Mark Allman’s Thoughts on reviewing
  - Timothy Roscoe’s Writing reviews for systems conferences
  - Register with HotCRP (find it in the course website)
Review form

- Overall merit
- Reviewer confidence
- Paper summary
  - Showing you understood the paper
- Paper strengths and weaknesses
  - Short sentences to the point
- Comments for authors
  - Details, avoid English/grammar/style (unless it gets offensive)
- Comments for PC
  - Concerns you may have, interpretation of your merit score (a discrete value), comparison with what you had in your pool, …
**Reviewing papers**

- **Goal**
  - Determine a paper’s suitability for a conference – WINE is very, very open so this is less important
  - Provide feedback to improve it

- **Things to ask**
  - Does the paper make a significant contribution, surprising results? Would the paper spark new research? Are the ideas clearly explained?

- **Consider the audience and the standards**

- **Process**
  - Read first for main ideas
  - Read it again and take notes
  - Start organizing your review
  - *(similar to Keshav’s approach)*
What it means to you

- **Reading and class discussion (10%)**
  - About 34 papers over the quarter (~4 papers per week), but you can “quickly” read most of them

- **Writing reviews and discussing papers as part of a program committee (20%)**
  - 7 reviews over the quarter

- **Presenting papers (20%)**
  - 2 paper presentations on each “side” of the midterm

- **Working on a project (50%)**
Giving a good research talk …

- Not easy – some guidelines
- The talk is just a taster, not an in-depth treatment of the work but intended to stimulate interest
  - Who is your primary audience? *Easy to do in our case*
  - If they remembers one thing from the talk, what should it be?
    - *In class, a bit more details to educate the rest of us*
- Think of what motivated you (or the authors)
  - Use examples to motivate the work and approach
  - Use examples to illustrate your points
- Saying enough without saying too much
  - Enough to convey your ideas, not to much to overwhelm your audience – follow a non-uniform approach
To include in your presentation

- Motivation
- Structured related work/background
  - What does the paper build on?
- Main ideas
- Evaluation approach
- Summary of key results
- Moving forward
  - Open issues
  - How has the paper been built on/cited/used?
Giving a good research talk

- Don’t put too much on the slide
  - Prune multiple times
  - People can only take very limited info per slide: 5+/−2 things
  - **Just one figure per slide!**
  - Don’t repeat what you plan to say
  - Use the slide header when possible

- Consider dropping the “overview/roadmap” slide

- Mind the time!

- …

- *Pick papers to present and email me*

- *We need 4 presenters for next week*
What it means to you

- **Reading and class discussion (10%)**
  - About 34 papers over the quarter (~4 papers per week), but you can “quickly” read most of them

- **Writing reviews and discussing papers as part of a program committee (20%)**
  - 7 reviews over the quarter

- **Presenting papers (20%)**
  - 2 paper presentations on each “side” of the midterm

- **Working on a project (50%)**
Project

- One single project – critical part of the course – 50%
- Your goal
  - Experiment with/measure an interesting system or property
  - Present your results at the end of the course
  - Write about it in a term paper
- Teams of 2+ people; based on topics you will be assigned a project leader (could be me, one of my students or somebody from outside the class)

- To keep in mind …
  - Paper registration / Submission:
    May 11, 18
From reading to research

- How can we get new ideas from studying papers describing solved problems?!
- Questions to ask
  - Unfinished business? Look for future work, directions, things not tried, open problems
  - Time – Revisiting old solutions in new, always different environments
  - Context – Good (and not so good) solutions in a different context (server replica selection from wired Internet to cellular networks)
  - Removing assumptions
Schedule

- Form a group – by Jan 9
- Project meeting with me – before Jan 12
- Project proposal – Jan 20
  - 2-pager; follow the website guidelines and post in Piazza
- Midterm presentation and report – Feb 19
- Final presentation – Finals week
- Final report due – Monday of finals week
  - HotNets format: submissions should contain six or fewer two-column pages …
    http://conferences.sigcomm.org/hotnets/2016/cfp.html
  - (look at https://www.overleaf.com/ )
Project presentations

• Midterm presentation and report – 4-6 slides
  – Project name & team members
  – Revised statement of project goals and list of new/interesting concepts to explore
  – List of issues addressed and pending
  – Updated project milestones, highlighting accomplishments to date, and schedule for the rest of the quarter

• Final presentation – ~10-15’ based on final report
Maybe counterintuitive, but don’t wait to write till the end

Abstract
  – What did you do, why is important & what are your high-level results?

Problem statement
  – What is the problem you tried to solve?

Prior work
  – How has the problem being dealt with before? Why was that not enough?

…
Final report

- ...
- Research approach
  - Approach to solving the problem? What did you design, built? What was your evaluation methodology?
- Results
  - How did you evaluate the work? What were your figures of merit?
  - (Not even this)
- Lessons learned and future work
  - What would you have done differently? What’s left for future work?
- Summary and conclusions
Summary

- You’ll read in depth about 7-12 papers
- Write reviews for 7 of them
- Skim++ another ~20
- Do 2-4 presentations
- Work in a project with a team

You’ll learn

- How to do experimentation at Internet scale
- State of the art and some open issues
- How to do reviews, present at a conference, discuss papers, carry on a research project …

Form a group if you haven’t already! Start reading