

Final Project

Due: June, 3rd 2021 11:59 PM (Pacific Time)

Now that you've gotten some practice reading and digesting research papers and reproducing a known result, it's time for the final project! This project will give you a chance to learn something new about an exciting piece of work, and will give you an introduction to the research process.

In the past, students have been able to recreate research, find flaws in published papers, and begun projects that eventually led to publication. We're excited to see what you do!

You may choose to either:

- **reproduce a result from a networking paper** that hasn't been reproduced in previous iterations of CS244, or
- **an original research project** in the area of computer networking

You may work on this project alone, or in groups of two. You can use [this thread](#) on Piazza to find a partner.

Much of this project is intentionally underspecified, and you should pick a project that you have sufficient time and experience to accomplish. Depending on what you propose, your project could require using a profiler, modifying the Linux kernel, implementing a custom simulator, learning a new scripting language, proving an interesting new result, or coordinating a software development effort across multiple group members.

You will likely need to read research papers outside of the class reading list or the particular paper you are trying to reproduce. (Check the references of any paper for the most closely related prior work, and read or skim the prior work!)

Reproduction Projects

For a **reproduction project**, here are some sample questions your project might try to answer:

- Does the primary result in the paper hold up? Your reproduction effort should be designed in a way that it can, in principle, detect if the original paper is wrong.
- Can you generate a plot or graph that looks the same as one of the main figures in the original paper, and explain where it came from?
- What kind of "reproduction" is appropriate? Any figure can be "reproduced" if you simply re-plot the authors' original data on the same axes. Can the data be generated afresh?
- Are the assumptions behind the data-generating process, and the way the system was compared with prior work, realistic? Do your results support the paper's claims about the generalizability, or applicability, of the results?
- What was difficult to reproduce? Are you able to package the contribution of the paper in a way that makes it easier to reproduce in the future?

- What happens if you vary a parameter the original experimenter didn't consider?
- Having reproduced the primary result, can you now extend or improve the work?
- Can you find a simplification to the system that still preserves most of the claimed improvements? If so, this can indicate that the original authors are mistaken about why their system outperforms prior work.

Selecting a paper

We would like you to reproduce a paper which has not yet been reproduced by a CS 244 student. The previous class projects are on the [Reproducing Network Research](#) blog.

You could reproduce an older, classic paper which hasn't been reproduced or one that was published more recently. To find papers, you could look at the programs of recent networking conferences, the websites of networking authors you know (e.g. ones we've read in class), papers which cite or are cited by networking papers, and so on.

Networking venues include:

- SIGCOMM: major conference in computer networking
- NSDI: networking and systems (overlap with OSDI)
- IMC: top-tier conference in Internet measurement
- CoNEXT: second-tier conference in computer networking
- OSDI: major conference in computer systems (sometimes including networking)
- SOSP: alternates years with OSDI
- USENIX ATC: second-tier conference in computer systems (sometimes including networking)
- HotNets: upcoming work and position papers in networking. Intended to be exciting, half-baked proposals.
- CCR: Computer Communications Review.

Original Projects

An **original project** can be a proof of concept of a novel networked system, an extension of existing work, or any research project that you could realistically extend to a conference-quality paper.

Note that original projects can be a bit more risky than reproduction projects, since research often does not go as planned.

Some ideas include:

- **New systems.** Questions to answer: how and why did you design the system the way you did? How does the system perform in comparison with prior work in similar areas? Why does it perform as it does (i.e., which design decisions account for what benefits)? What are the system's weaknesses and limitations?

- **Measurements.** What can you learn (about the Internet, about another network, about applications, about people, about spam email, about cryptocurrency, about online video games, about social science) by using the tools of computer networking? What kind of measurement tools or techniques did you use, and what did they tell you about something in the real world? How can somebody reproduce your measurement? What are some weaknesses or potential confounders of your measurement?
- **Any other creative project** using the tools or problem domain of computer networking.

Timeline

Project Proposal

Due Tuesday, May 4th 2021, 11:59pm

Before you submit your proposal, we encourage you to talk to us (e.g. in office hours, or in lecture).

Please propose the topic for your project. If you're not sure of what you want to do by the proposal deadline, you can submit a few possible options and we can help you decide.

Please be wise and mindful of your (other) deadlines and other commitments when picking the project.

Please submit one page or less on canvas, which includes:

- Who you will be working with (we encourage groups of two, but working alone is okay)
- The particular results you would like to replicate, or the overall goal of your original project.
- A brief outline of incremental steps you will need to do to finish the project as well as a timeline. The goal here is to convince both us and yourself that your project is neither too small nor too big.

Weekly Status Updates

Please send a **weekly email** the course staff which includes:

1. What you did this week
2. What you need to do next week to stay on track.
3. If there's anything the course staff can help with

These emails will not be graded; the idea is to make sure you are making incremental progress and let us help you with small issues before they become major problems.

We also encourage you to visit the office hours to discuss your project as your project evolves.

Intermediate project report

Due Tuesday, May 18 2021, 11:59pm

Please include the following in the report:

- A draft introduction for the final report.
- A baseline reproduction result. This doesn't need to be correct, but we would like to see progress
- The status of the project. What's working, what are you having trouble with?
- If you've reached out to the original paper authors, any responses you've gotten.
- Plan for the remaining time

Formatting:

Please use the [SIGCOMM LaTeX template](#). If you would like to use something else besides LaTeX, please make sure your paper adheres to the [SIGCOMM style guidelines](#)

If there is anything you need from the course staff, please let us know.

Writing an Introduction

The introduction is the first, and unfortunately sometimes only, part of your research that someone will read. It's important for it to get the main points of your result across in a clear and precise way.

Please make sure your introduction does the following:

- States what problem the original paper was solving, and why the original paper was important. This should be self-contained: a reader shouldn't have to read the original paper to understand your introduction.
- Includes the result of the original paper
- *Briefly* describe your reproduction methodology (in 1-3 sentences)
- States the high-level results of your reproduction (when you have them)

For examples, you could refer to:

- The introductions to networking papers we've read so far.
- The introductions from past reproductions on the [Reproducing Network Research](#) blog

Mid-quarter check-in

During the week of May 17th, you'll have **one mandatory meeting** with the course staff for around 15-20 minutes. The goal is to make sure you're on track to complete the assignment.

Presentations

Held on Tuesday, June 1 2021

More info to come.

Final Report

Due Thursday, June 3 2021, 11:59pm

Your final deliverable will be a **4-6 page report** about your project. You will be expected to:

- Introduce the project: place it in context, give an overview, and summarize your major contributions or findings
- Discuss prior work and the landscape of the area
- Explain your methods and/or the design of your system
- Present the experimental setup in detail (assuming there are experiments, rather than proofs)
- Describe the results in figures and text
- Discuss the meaning of the results and what the reader should take away
- Describe the limitations and weaknesses of your system, measurement, or effort at replication
- Propose a plan for future work in the area. What would you do next if you had more time?

For example reproductions, please see the [Reproducing Network Research](#) blog. We are planning on posting your final writeup on the blog for posterity, but please email us in advance if you would like us to *not* post your writeup.

Grading

You'll be graded out of 40 points total.

The grades will be based on the following:

- Intermediate progress report
- Final report
- Quality of experimental setup and results
- If you end up doing a presentation, this will only boost your grade.