

CS 6290

Privacy-enhancing Technologies

Department of Computer Science

Slides credit in part from D. Song

Tutorial 5

Yufei CHEN

CS Department
City University of Hong Kong

About the Group Project?

- Project outcomes should include:
 - report + oral presentation (MUST)
 - implementation code (for project involve programming)
- Group size: 3~4 members.
 - Individual project is OK, but it is **very challenging**.

Class Project Categories

- Systematization of Knowledge
- Measurement/empirical study
- New design and implementation

Systematization of Knowledge (SoK)

- Goal:
 - Survey work in an area/on a topic
 - Establish a framework & extract insight
 - Conduct analysis and experiments/measurements as needed (extensive analysis and experiments)

Systematization of Knowledge (SoK)

Example SoKs:

- [SoK: Decentralized Finance \(DeFi\) Attacks \(IEEE S&P '23\)](#)
- [SoK: How Robust is Image Classification Deep Neural Network Watermarking? \(IEEE S&P '22\)](#)
- [SoK: Privacy-Preserving Data Synthesis \(IEEE S&P '24\)](#)

(More SoKs can be found from  <https://oaklandsok.github.io>)

Evaluation:

- Does it cover representative works in the area/on the topic?
- What are the framework & insights?
- Are analysis and/or experiments sufficient in supporting the insights?

Measurement/Empirical Study

Goal:

- To establish a rigorous and repeatable methodology for studying a system or phenomenon using empirical data, e.g., developing a benchmark suite for evaluating the performance of different blockchain consensus algorithms, or creating a repeatable methodology for measuring the effectiveness of various online tracking prevention techniques.
- Study different aspects:
 - Incentive structures, risks, stabilities, considering incentive alignment
 - Throughput, latency, security, etc. from a systems perspective

Measurement/Empirical Study

Methodology:

- Gather data
- Identify key metrics and questions for measurement
- Analyze data
- Extract insights

Measurement/Empirical Study

Project evaluation:

- What are the key metrics and questions for measurement?
- Is the data sufficient to measure the key metrics & answer the questions?
- What are the insights?
- Is the analysis repeatable?

Measurement/Empirical Study

Sample project: DeFi Attacks Empirical Study

- Sample Paper: <https://arxiv.org/abs/2003.03810>
- Sample Paper: <https://arxiv.org/pdf/2106.06389.pdf>
- Possible breakdown
 - Task 1 - Select a list of related attacks (e.g., on-chain price oracle manipulation)
 - Task 2 - Reproduce these attacks by forking the blockchain and find optimisations to find the optimal attack vector.
 - Task 3 - Perform additional analysis to discover new findings that have not been made public via social media or articles.

Measurement/Empirical Study

Another sample: rethinking previous PETs

Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models

Shawn Shan, Jenna Cryan, Emily Wenger, Haitao Zheng, Rana Hanocka, Ben Y. Zhao

Department of Computer Science, University of Chicago

{shawnshan, jennacryan, ewillson, htzheng, ranahanocka, ravenben}@cs.uchicago.edu

Glaze: Protecting Artists from Style Mimicry by Text-to-Image Models. USENIX Security 2023.
(Distinguished paper award 🏆)

Measurement/Empirical Study

Another sample: rethinking previous PETs

**ADVERSARIAL PERTURBATIONS CANNOT RELIABLY
PROTECT ARTISTS FROM GENERATIVE AI**

Robert Hönig
ETH Zurich

Javier Rando
ETH Zurich

Nicholas Carlini
Google DeepMind

Florian Tramèr
ETH Zurich

Measurement/Empirical Study

Another sample: rethinking previous PETs

ADVERSARIAL PERTURBATIONS CANNOT RELIABLY PROTECT ARTISTS

Robert Hönig
ETH Zurich

Javier R
ETH Zu

ABSTRACT

Artists are increasingly concerned about advancements in image generation models that can closely replicate their unique artistic styles. In response, several protection tools against style mimicry have been developed that incorporate small adversarial perturbations into artworks published online. In this work, we evaluate the effectiveness of popular protections—with millions of downloads—and show they only provide a false sense of security. We find that low-effort and “off-the-shelf” techniques, such as image upscaling, are sufficient to create robust mimicry methods that significantly degrade existing protections. Through a user study, we demonstrate that all existing protections can be easily bypassed, leaving artists vulnerable to style mimicry. We caution that tools based on adversarial perturbations cannot reliably protect artists from the misuse of generative AI, and urge the development of alternative protective solutions.

New Design and Implementation

Project evaluation:

- Is the problem clearly defined?
- What is the new approach/solution?
- Do the experiments properly evaluate the solution? How well does the solution improve over previous solutions?
- Would this project be a good workshop or (possibly with some additional work) conference paper?

New Design and Implementation

Sample project ideas:

- New approach for decentralized identity
- New design for privacy-preserving financial services
- New interoperability solutions
- New zero-knowledge proof applications
- Innovative decentralized systems like new execution environments etc.

Project Topic Examples

Cryptocurrency & Blockchain Systems

- * Consensus Protocol Analysis & Comparison
- * Smart Contract Vulnerability Analysis
- * Blockchain Scalability Solutions
- * Cross-Chain Interoperability Mechanisms
- * Alternative Consensus Mechanisms

Data Privacy

- * Differential Privacy for Data Publishing
- * Privacy Attacks on Machine Learning
- * Anonymization Techniques for Datasets
- * Survey of Privacy-Enhancing Technologies (PETs)
- * Legal and Ethical Aspects of Data Privacy

Confidential Computing Technologies

- * Zero-Knowledge Proof Applications
- * Multi-Party Computation for Secure Analytics
- * Secure Hardware for Blockchain Applications
- * Privacy-Preserving Smart Contracts
- * ZKP for Identity and Authentication

Privacy in the Wild

- * Online Tracking Measurement & Analysis
- * Private Machine Learning Techniques
- * Browser Privacy Extension Evaluation
- * Privacy Implications of IoT Devices
- * User Perceptions of Online Privacy

Concerns?

Limited time to fully investigate a new era?

Technical complexity overwhelming?

Teamwork troubles?

Access to resources and data?

Fear of the unknown and making mistakes?

...

Concerns?

Limited time to fully investigate a new era?

- Projects are designed to be *focused* explorations, not exhaustive research.
- Consider to start from the field you are familiar with.
- Seek help from the TA.
- Teamwork!

Concerns?

Technical complexity overwhelming?

- Start simple, build up
- Collaboration is key

Concerns?

Teamwork troubles?

- Early Planning & Clear Roles
- Regular Team Meetings (even brief check-ins)

Concerns?

Access to resources and data?

- The scale of the experiment is not our primary concern (e.g., when you work on machine-learning-related work, consider to use small models.)
- Try to select topics where public data are available

Concerns?

Fear of the unknown and making mistakes?

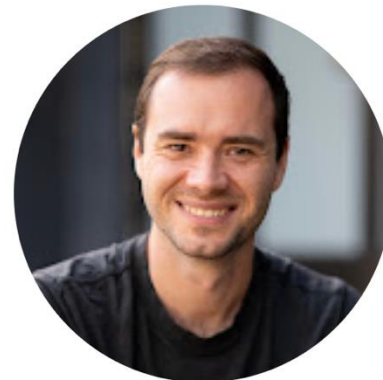
- Mistakes are valuable for growth.
- Grading will consider effort, learning, and progress, not just “perfect” outcomes.

One more suggestion

- AI can also be your teammate!
 - Intelligent Search & Literature Review
 - Idea Generation & Refinement
 - Rapid Information Synthesis & Summarization
 - Basic Code Prototyping (if applicable)
 - Writing & Communication Enhancement

How I use LLMs

<https://www.youtube.com/watch?v=EWvNQjAaOHw>



Andrej Karpathy

@AndrejKarpathy · 696K subscribers · 18 videos

More about this channel ...more

x.com/karpathy and 3 more links



Subscribed



One more suggestion

- AI can also be your teammate!

WARNINGS


- Comply with the CityU's policy about the usage of GenAI
- AI is not a fact source
- Responsible use

AI isn't replacing *you*, it's *empowering* you to learn faster, research smarter, and achieve more in the limited project timeframe.



Policy of GenAI

- 1) Students are **not allowed** to use GenAI for programming tasks
- 2) When writing assignments and reports, you can use GenAI to help you do search and improve the writing quality. (Remember to acknowledge the usage of AI)

Group Project Proposal (Week 7)

- TA will release **submission instructions**
-  **The proposal Must Include:**
 1. **Problem Motivation:** Why does your topic matter?
 2. **Difficulty:** Challenges and how you'll address them.
 3. **Timetable Plan:** Weekly milestones + task ownership.

Group Project Proposal (Week 7)

-  **Scoring Criteria:**
 - Relevance/Importance of problem, feasibility, timeline realism.
-  **Why It Matters:**
 - Top-scoring teams pick **final presentation timeslots first**
 - Refine project direction early.
 - Gain actionable feedback from instructors/TAs.
- Scores do NOT impact coursework grades—**focus on feedback!**