Outsourcing of data storage

 Txn1

 Txn2

 [Diagram showing data storage and transactions]

 2
Why is privacy of outsourced data important?

Google settles $5 billion consumer privacy lawsuit
Goal of this class:

Learn to design data systems that protect data privacy!
Challenges with privacy preserving data systems

Encrypt the data!

- What do we lose??
  **Functionality!!**

Plaintext database allows you to query on Id, Name, Age (e.g., names of all users < 50)

Cannot trivially support such queries on encrypted data!
In this course..

• We will first study **CryptDB**: an encrypted database that supports many type of queries even when data is encrypted

• How? → Using different types of encryption mechanisms. Primer in the next lecture!

• Is encrypting the data always sufficient to protect privacy??

• No!! Learn about **inference attacks** (on CryptDB)!!
Motivated by inference attacks, we will study systems that protect against such attacks using 4 different techniques

1. Oblivious RAM and alternatives
2. Secure multi party computation (using garbled circuits and secret sharing)
3. Trusted hardware enclaves
4. Private information retrieval

I will introduce the technique, which will be followed by student presentations of relevant papers
Other challenges with privacy preserving data systems

• Apart from supporting complex types of queries, plaintext databases are scalable, tolerate unexpected crash or network failures, and provide high performance by allowing concurrent accesses.

• Can encrypted databases give similar guarantees trivially?

• No!! We will learn about privacy leakages due to concurrency, scalability, and fault tolerance. And how different systems hide these leakages.

• This is primarily a data systems course!
Logistics

• Class is Tuesdays and Thursdays, 10:00 AM to 11:20 AM
  • Class may be virtual depending on the weather; you will have at least 15 hr notice

• Course website: https://cs.uwaterloo.ca/~smaiyya/cs848/

• I will introduce the necessary cryptographic techniques, which will be followed by student presentations of relevant papers

• Email me (smaiyya@uwaterloo.ca) for any questions and to request office hours
  • prefix the email with [CS848] for a timely reply

• We will use Piazza for class announcements. Please sign up using this link: https://piazza.com/uwaterloo.ca/winter2024/cs848002
Course Components
(may vary slightly depending on the class size)

• Paper reviews – 20%
• Paper presentation – 15%
• Class participation – 15%
• Course project – 50%
Paper reviews

• Read and write a review for two papers per week. Due at 1PM the day before the section!

• 500 words limit and contain the following sections:
  1. A concise summary of the paper (1 paragraph)
  2. A list of the paper's main strengths (at least 2 bullet points)
  3. A list of opportunities for improvement (at least 2 bullet points)
  4. Critical analysis and comments

• All reviews will be made public (anonymously) by 2PM
Paper reviews (cont.)

• Review grading
  • Complete (2 points): adheres to the reviewing guidelines (last slide), clearly demonstrates that the reviewer has read and thought about the paper
  • Partially Complete (1 point): Misses some but not all the reviewing guidelines, demonstrates that the reviewer has some understanding of the paper
  • Incomplete (0 points)

• We will use HotCRP for reviews: https://uwaterloo-cs848w24.hotcrp.com/
Paper presentation

• Each assigned paper has a primary and a secondary student

• Primary student creates a presentation and presents the paper (see course website for what to include in the presentation)

• Secondary student will read all the reviews made public at 2PM the afternoon before the section

• Both primary and secondary will lead the discussion. Each student must be a primary for 1-2 papers and secondary for 1-2 papers (depending on class size)

• Note: primary need not write the review

• Sign up link: https://docs.google.com/spreadsheets/d/1fNs-sfJH8HuSkdCuvr7rWjTQAHt7YwPAtdUs2ZyJNk/edit?usp=sharing
Project

• Done individually or in groups
• Original research projects related to data systems with privacy guarantees
  • If in doubt, please talk to me at the earliest

• Three deliverables
  • Proposal – due Feb 4th
    Details: One page with problem statement, context and motivation, and a high-level overview of related work
  • Final presentation – due March 28\textsuperscript{nd} or April 2\textsuperscript{nd} (will be announced)
  • Final report – due 1 week after the presentation
    Details: 6-page conference-style paper with problem statement and motivation, design, evaluation, related work, and future research directions
Final remarks

• This is a grad course – ‘you reap what you sow’

• Please be active in class discussions – it makes sections interesting for everyone! But please be respectful & mindful of others.

• There are no ‘bad’ or ‘stupid’ questions or ideas! If you are hesitant to open-up publicly, please reach out and I will provide a safe space.

• Finally, let’s learn from each other and have fun in the process!