Privacy-Preserving Email Search

Rene Reyes (MIT), Christophe Hauser (ISI)

Problem Statement

- Popular email services (e.g.: GMail) store content of emails in servers that can be compromised
- End-to-end encryption provides privacy, but no search functionality
  - This requires emails to be stored locally to allow for full-text search
- We explore the use of Symmetric Searchable Encryption (SSE), a tool that has been widely studied in theoretical cryptography, to add search functionality

Background: SSE

- Secure search index based on keywords
- Static and **Dynamic** constructions
- Privacy guarantees:
  - Forward Privacy: Updates
  - Backward Privacy: Deletions
  - Hiding Access Pattern: Documents matching query
  - Hiding Search Pattern: Correlate different queries
- Constructions/Countermeasures:
  - Forward + Backward Privacy: Oblivious Dynamic Cross-Tags (ODXT)[1]
  - Access Pattern: Redundant Encoding + Noise[2] provide Differential Privacy (DP)
  - Search Pattern: Group-Based Construction (GBC)[3] for keywords
- Privacy-Performance tradeoffs: Computation time, server-side storage, local storage

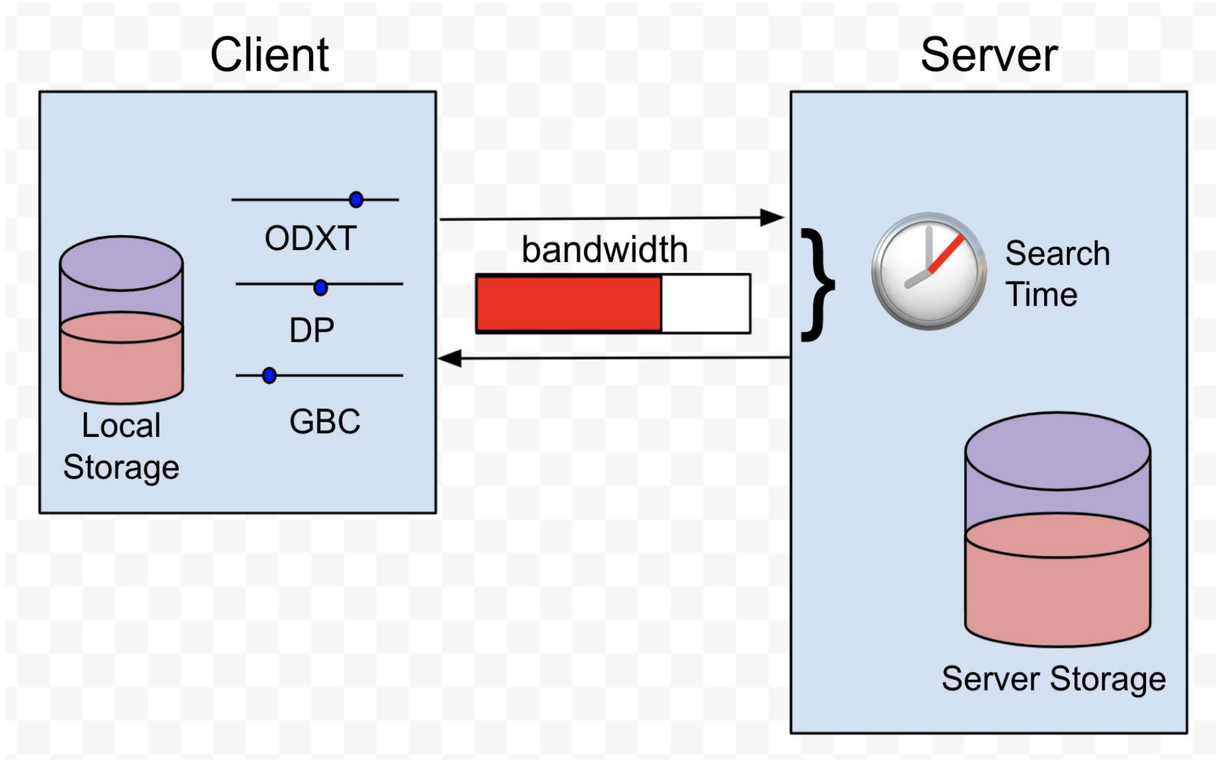
Threat Model + Privacy Guarantees

- *Honest-but-curious* server: follows protocols but may try to learn private information
- Server can also *inject files* into search index by sending emails to the user
- Give users control of privacy guarantees:
  - Provide an understandable explanation of privacy threats as a result of parameter settings
  - Users can make decisions based on desired performance and privacy protection
  - System can adjust options given to user based on storage and bandwidth constraints

Design + Implementation

- Our system is compatible with existing end-to-end encrypted email services
  - Search index is updated by local email client after decrypting received emails
- Python ODXT implementation + existing open-source tools with python wrappers
- Adapted use of DP countermeasure for the dynamic setting

Evaluation Plan + Future Work



- Privacy evaluation is theoretical and based on building blocks
- Performance Exploration:
  - Different security parameter settings
  - Measure various performance metrics (shown to the right)
  - Use results for options given to users in privacy setup
- Potential user study:
  - Receive feedback on interface
  - How well privacy setup communicated tradeoffs with performance

1. Sikhar Patranabis and Debdeep Mukhopadhyay. Forward and backward private conjunctive searchable symmetric encryption. Cryptology ePrint Archive, Report 2020/1342, 2020. https://eprint.iacr.org/2020/1342
2. Guoxing Chen, T. Lai, M. Reiter, and Yinqian Zhang. Differentially private access patterns for searchable symmetric encryption. IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, pages 810–818, 2018.
3. Chang Liu, Liehuang Zhu, Mingzhong Wang, and Yuan Tan. Search pattern leakage in searchable encryption: Attacks and new construction. Cryptology ePrint Archive, Report 2013/163, 2013. https://eprint.iacr.org/2013/163.

If interested contact Rene David Reyes: rdreyes@mit.edu
Work performed under REU Site program
supported by NSF grant #2051101

USC Viterbi
School of Engineering
Information Sciences Institute