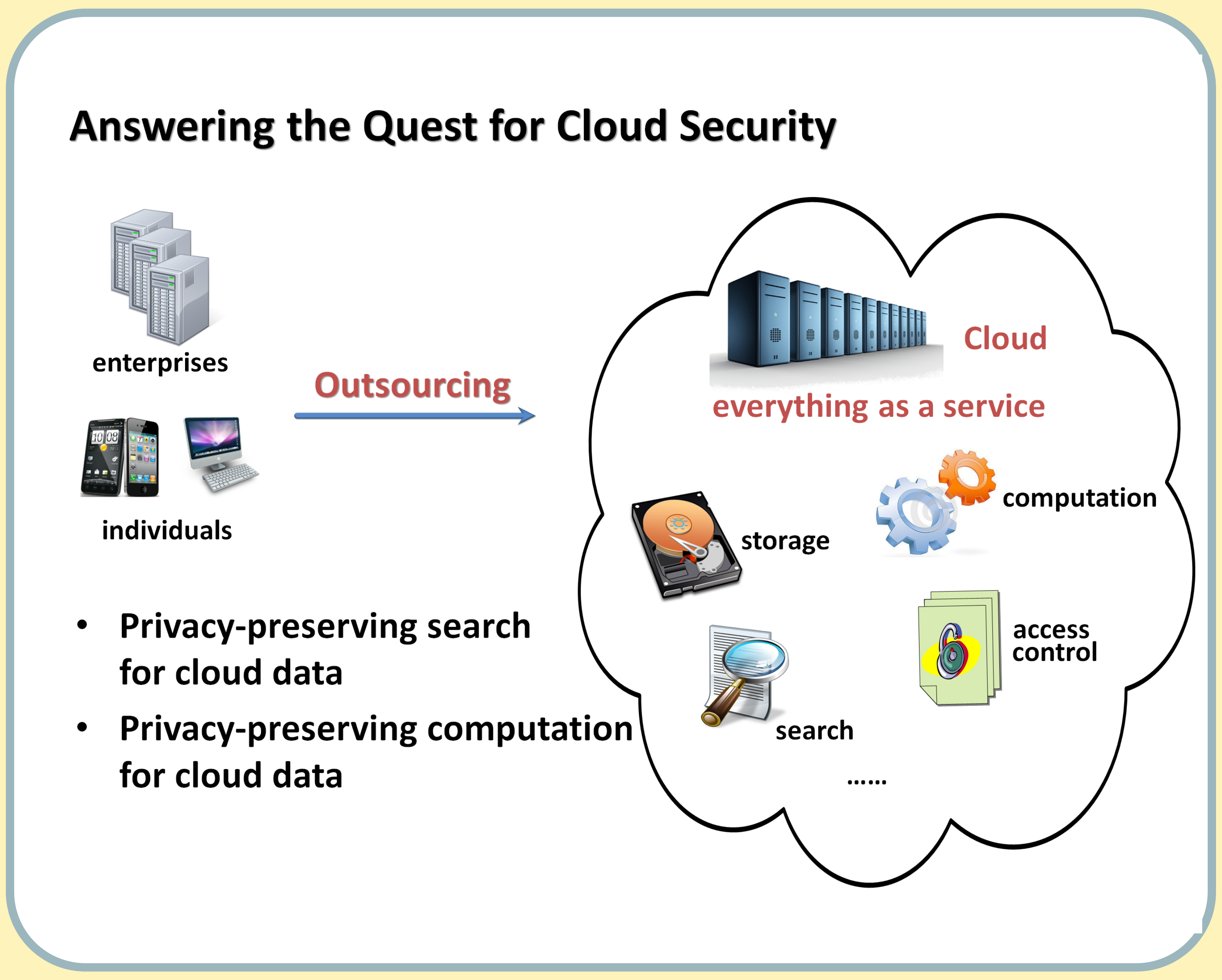
Privacy-preserving Search and Computation for Cloud Data

Prof. Kui Ren Dept. of CSE State University of New York at Buffalo

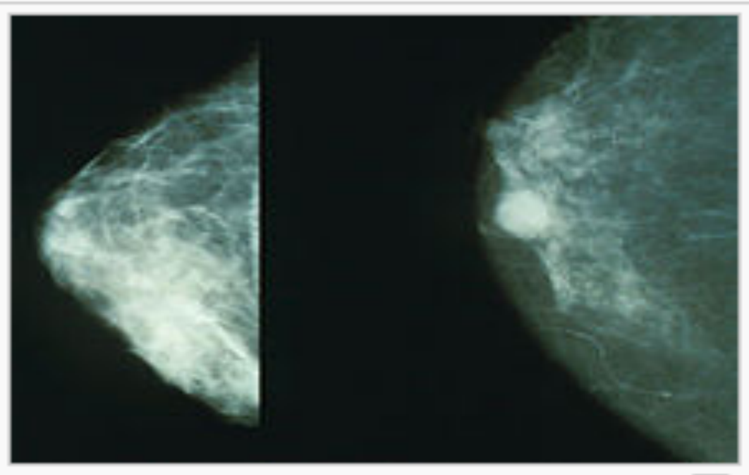
Answering the quest for cloud data security

Cloud computing serves as natural hub hosting massive data continuously generated by the Internet and social media, which take various different forms, e.g., text, picture, multimedia, etc. Numerous cloud services are being deployed adopting such a model. While the merits of cloud services can be easily perceived, their security and privacy risks still largely remain a challenge.



Answering the Quest for Cloud Security

- Privacy-preserving search for cloud data
- Privacy-preserving computation for cloud data



Normal (left) versus cancerous (right) mammography image.

Remote Sensing of Washington Area

Modality	One Image (bits)	# of image/exam	Size for 1 exam
Digital Mammography	**4000 x 5000** x 12 (2 bytes)	4	**160 MB**

Data encryption is a must for data confidentiality, but it also necessitates the need for developing effective searching techniques over encrypted cloud data of massive scale.

Cloud data are also being frequently processed for the data mining purpose. It is highly critical to develop privacy-preserving and proof-carrying computation and data mining mechanisms that suit for large-scale applications.

Challenges and Approaches

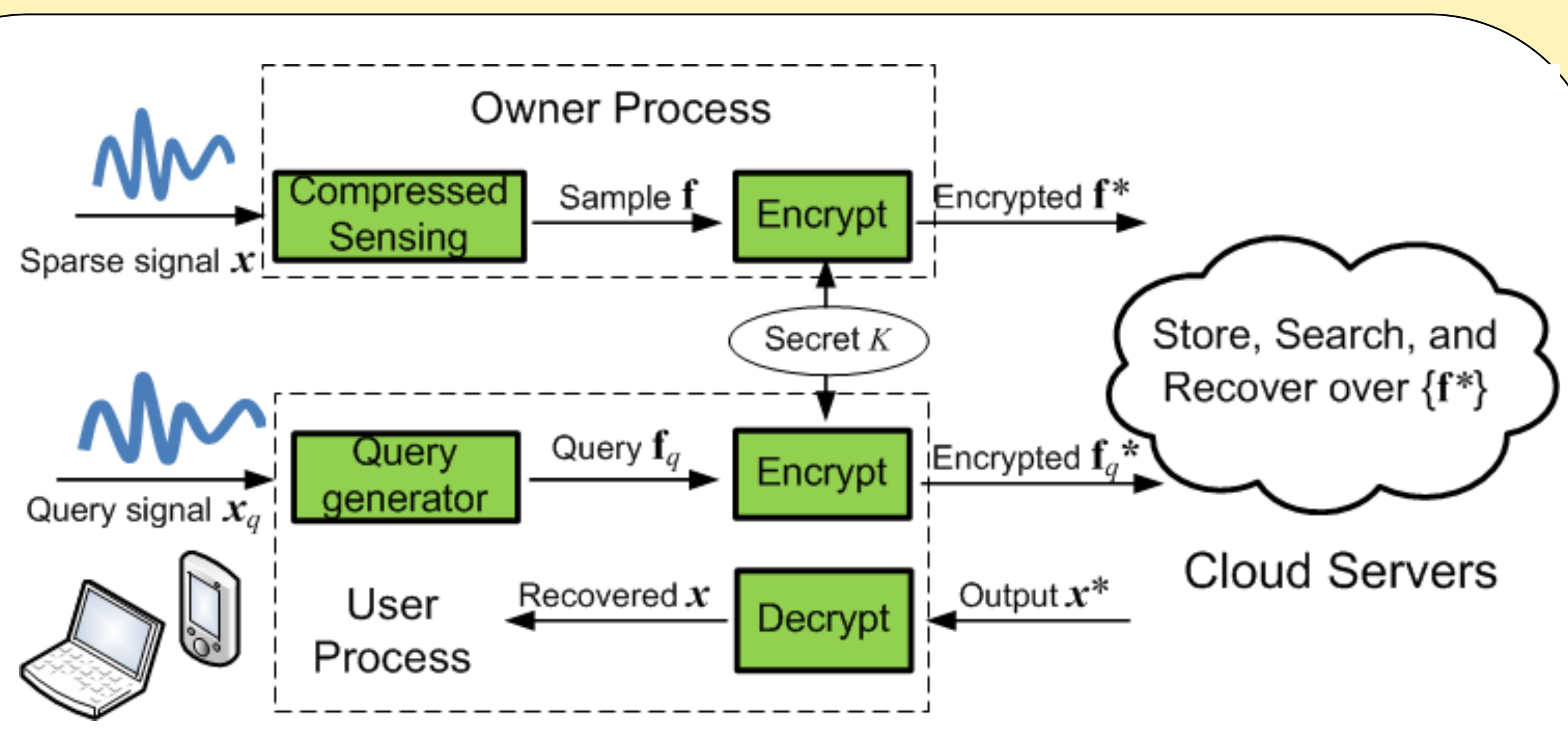
Privacy-preserving Data Search

- Most existing searchable encryption techniques support only simple keyword/predicate matching functions. But they only support simple text data, and is with very limited in functionality, usability, scalability, and performance.
- Expertise from different communities including cryptography, security, database, information retrieval, algorithms, and distributed systems, needs to join together to solve the challenge.

Secure Computation Outsourcing

•Theoretically, we can rely on fully homomorphic encryption (FHE), to construct a universal solution that is perfectly secure. The performance of FHE is however totally unacceptable as of today or in the near future.

•Our approach is to understand the nature of an application and its security requirements and develop application-specific solutions that are highly customized and achieve desirable trade-offs among privacy protection, performance, and other factors.

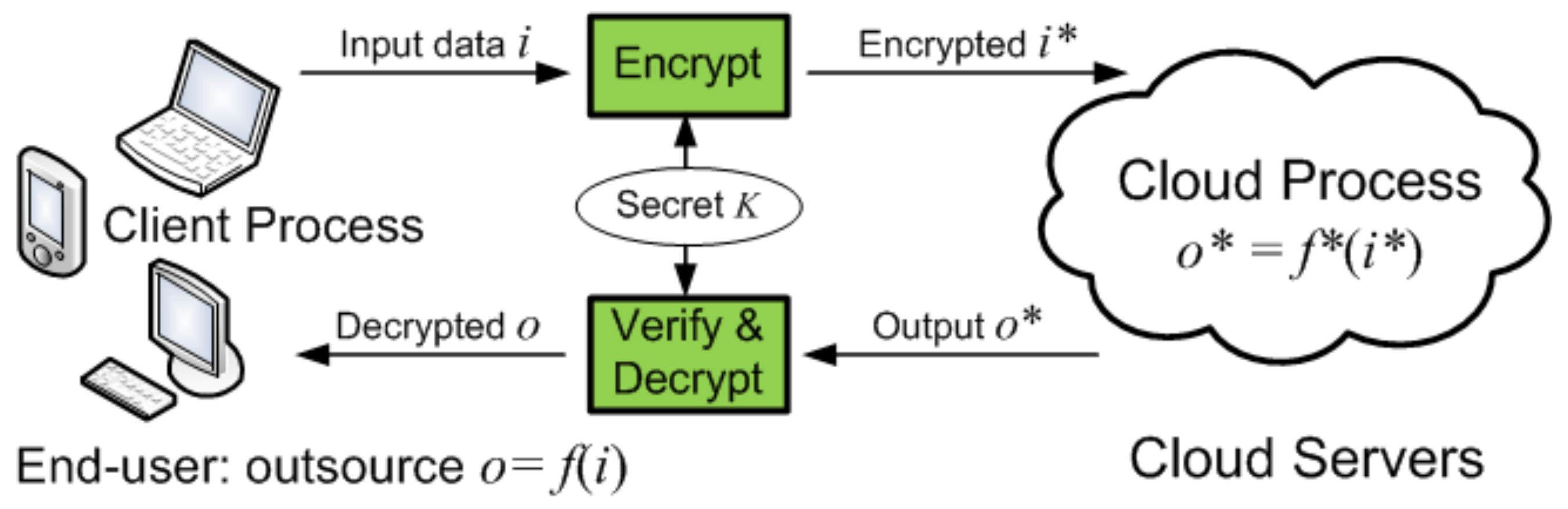


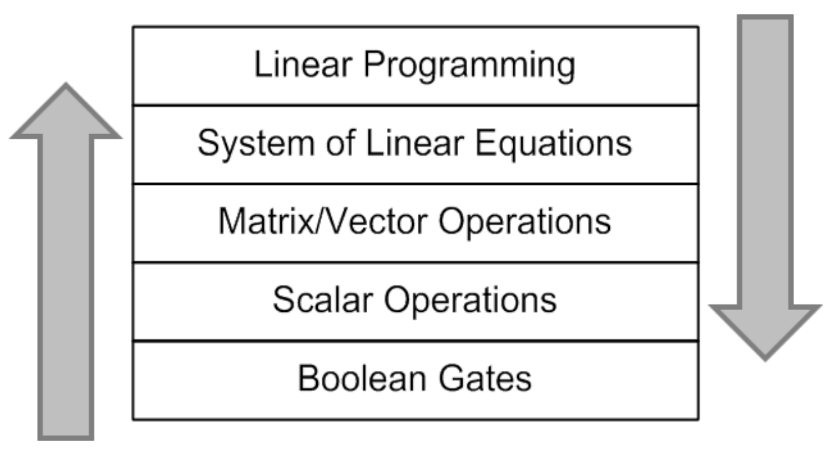
Current focus: High-dimensional data recovery and search in the context of compressive sensing; Content-based image retrieval

Key observation: Compressed sample data can serve for dual purposes: image recovery and content-based image retrieval.

Key techniques: Secure local/global feature based image retrieval

Secure searchable index leveraging locality sensitive hashing.





Current focus:

Systematically exploit security/efficiency tradeoffs, by interpreting computations as operations at different abstraction levels organized in a hierarchy.

Exemplary application include linear programming and linear equation.

$$\begin{array}{ll}\text{minimize} & \mathbf{c}^T\mathbf{x}\\ \text{subject to} & \mathbf{Ax} = \mathbf{b}\\ & \mathbf{Bx} \ge \mathbf{0}\end{array} \quad \xleftrightarrow[(\mathbf{Q}, \mathbf{M}, \mathbf{r}, \boldsymbol{\lambda}, \gamma)]{} \quad \begin{array}{ll}\text{minimize} & \mathbf{c'}^T\mathbf{y}\\ \text{subject to} & \mathbf{A'y} = \mathbf{b'}\\ & \mathbf{B'y} \ge \mathbf{0}\end{array}$$

Interested in meeting the PIs? Attach post-it note below!

National Science Foundation
WHERE DISCOVERIES BEGIN

The 3rd NSF Secure and Trustworthy Cyberspace Principal Investigator Meeting
January 9-11, 2017
Arlington, Virginia

University at Buffalo
The State University of New York