Searchable Encryption

Douglas Choi
CECS 579

Motivation

- ➔ growing demand for storage of confidential data
- ➔ use of third-party cloud storage solutions
- ➔ we primarily access data by search
- ➔ ***How do we search on encrypted data?***

Motivation

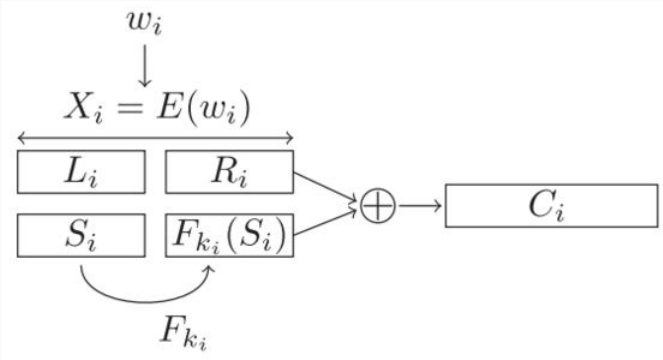

Encrypted documents
Encrypted keywords
Client
Search token
Server
Eve
Matching documents

Adversary Model

- ➔ Adversary can be a server or database administrator
- ➔ Security against adaptive chosen keyword attack - “CKA2”
  - ◆ Adversary cannot determine the contents of the documents or content of keyword…
    - even if adversary observes document ciphertext and search results
    - even if keywords are chosen by the adversary
    - even if keywords are chosen based on previous search history

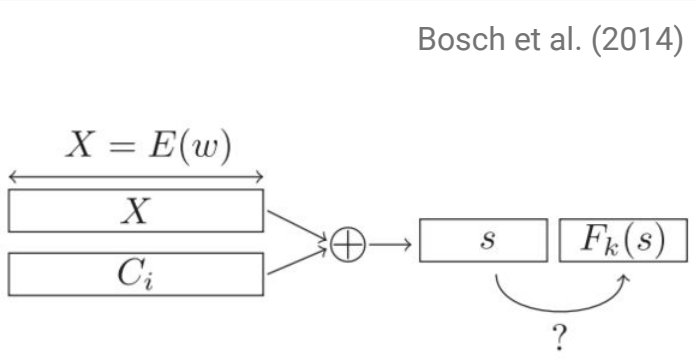
Searchable Symmetric Encryption (SSE)
Song, Wagner, Perrig (SWP)

Encrypt



Search

Bosch et al. (2014)

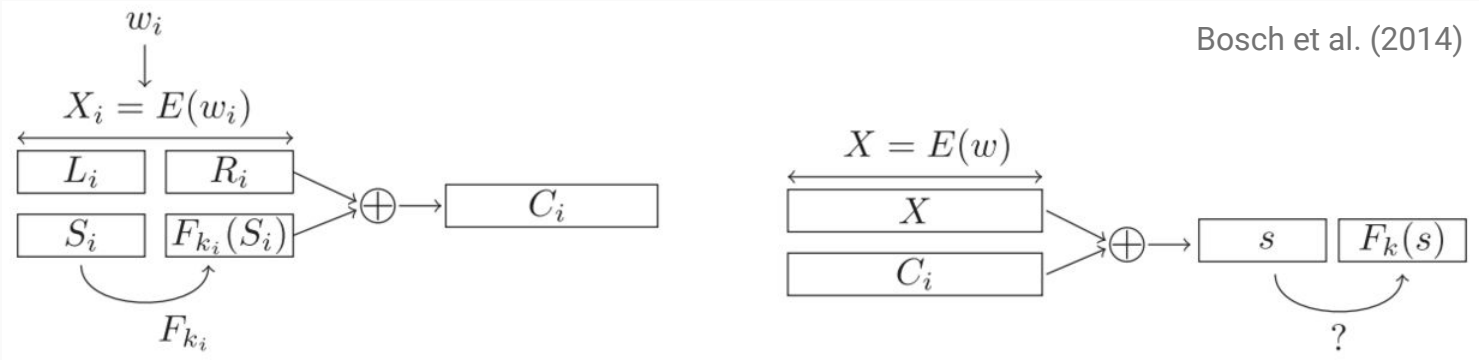


w	keyword
$X_i = E(w_i)$	deterministic encryption of w_i
S_i	pseudorandom value
k_i	key derived using a PRF of L_i
$F_{k_i}(S_i)$	hash of S_i using key k_i
C_i	ciphertext of w_i

Searchable Symmetric Encryption (SSE)
Song, Wagner, Perrig (SWP)

Encrypt

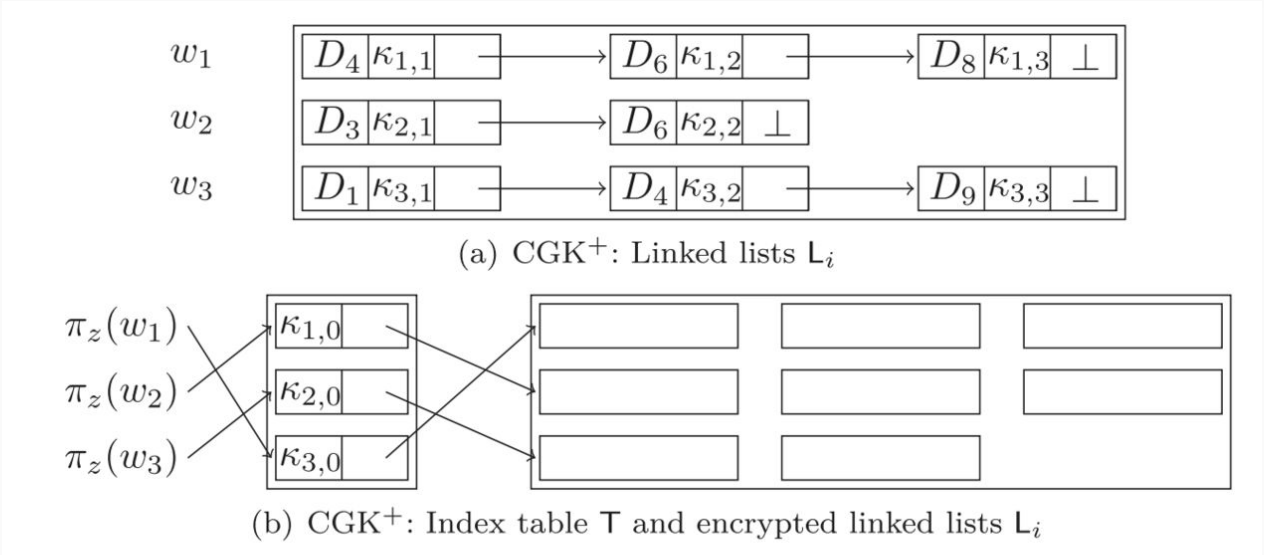
Search



➔ Slow search time: *O(documents x words)*

➔ This is not CKA2 secure because can learn the words using frequency analysis.

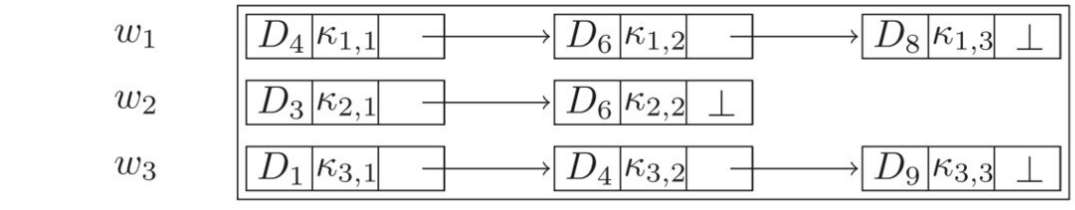
Searchable Symmetric Encryption (SSE)
Curtmola, Garay, Kamara (CGK)



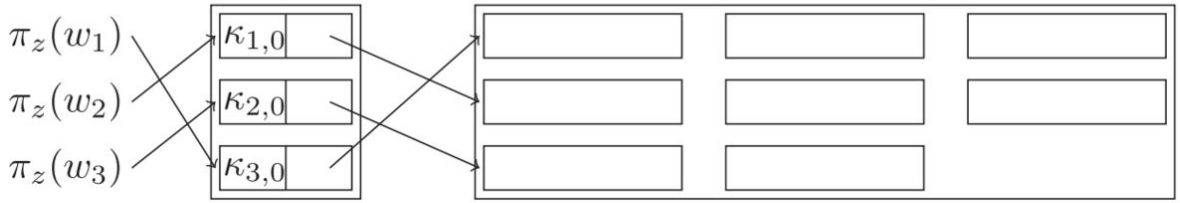
- ➔ Inverted index
- ➔ Consists of a linked list per distinct keyword
- ➔ Each node contains the document id and the key used to encrypt the next node
- ➔ Lookup table that maps the value of a PRF with some key z of the keyword to the head node

Bosch et al. (2014)

Searchable Symmetric Encryption (SSE)
Curtmola, Garay, Kamara (CGK)



(a) CGK$^+$: Linked lists L_i



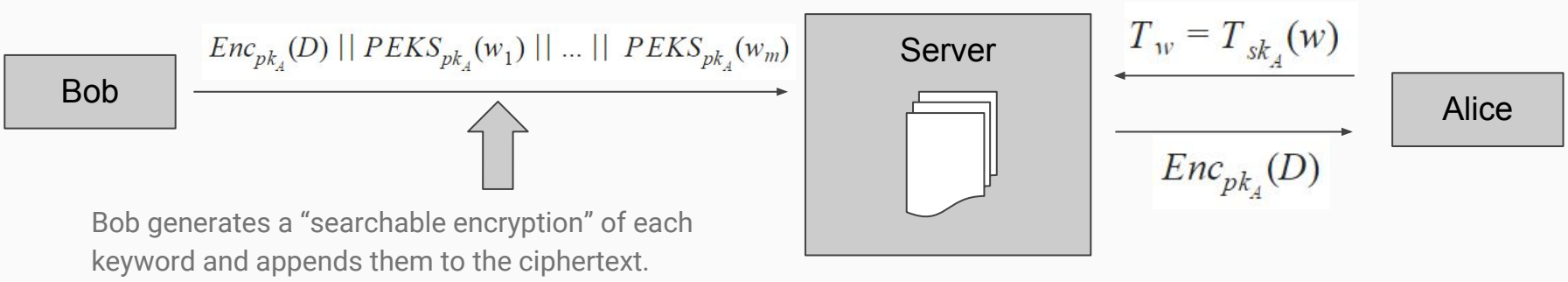
(b) CGK$^+$: Index table T and encrypted linked lists L_i

➔ High overhead

➔ Difficult to update documents

Bosch et al. (2014)

Public-key encryption with keyword search (PEKS)
Boneh, Crescenzo, Ostrovsky (BCO)



Bob generates a “searchable encryption” of each keyword and appends them to the ciphertext.

Comparisons

scheme	security	search time
SWP	CPA	O(mn)
CGK	CKA	O(r)
BCO	CKA	O(nv)

n = number of documents
m = number of keywords per document
r = number of documents containing keyword w
v = number of distinct keywords per document

Current Research

- ➔ Many different (and very complex) schemes under active research
- ➔ How to extend this to multiple writers and multiple readers
- ➔ How to improve search time and index sizes
- ➔ How to query for more complex search queries
- ➔ Schemes still reveal access patterns and search patterns

References

1. D.X. Song, D. Wagner, and A. Perrig. *Practical Techniques for Searches on Encrypted Data*. In 2000 Proceedings of the 2000 IEEE Symposium on Security and Privacy. 2000.

2. R. Curtmola, J. Garay, S. Kamara, and R. Ostrovsky. *Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions*. In 13th ACM Conference on Computer and Communications Security Proceedings, pages 79 - 88. October, 2006.

3. Dan Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano. *Public Key Encryption with Keyword Search*. In Eurocrypt 2004, LNCS 3027, pages 506-522, 2004.

4. C. Bosch, P. Hartel, W. Jonker, and A. Peter. *A Survey of Provably Secure Searchable Encryption*. In ACM Computing Surveys, Vol. 47, No. 2, Article 18. August, 2014.