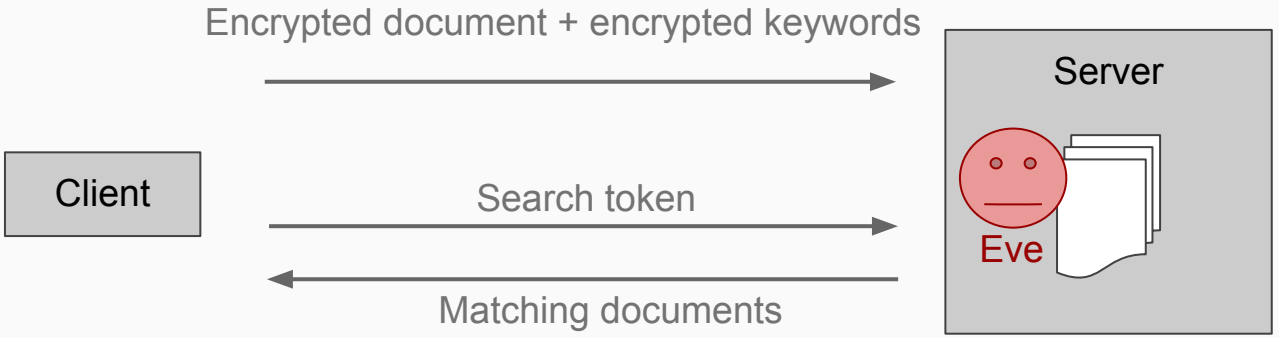
Searchable Encryption

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 securely?***

Motivation

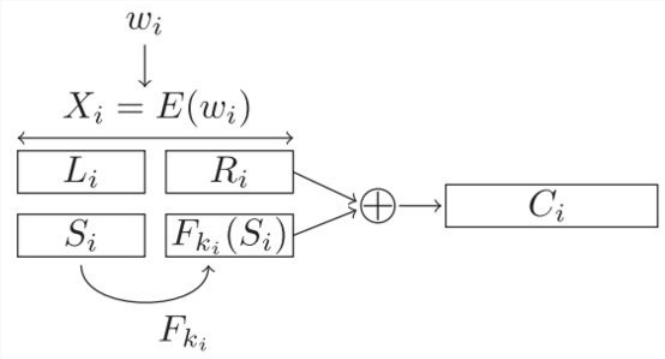

Encrypted document + encrypted keywords
Server
Client
Search token
Eve
Matching documents

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 contents of keywords…
    - even if adversary observes document ciphertext, search tokens, and search results
    - even if adversary can keep a history of search tokens and search results

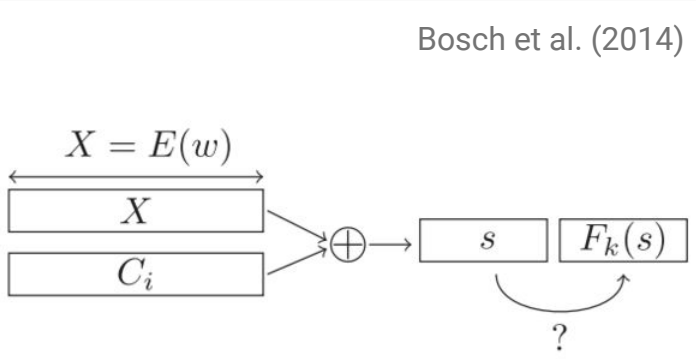
Searchable Symmetric Encryption (SSE)
Song, Wagner, Perrig (SWP)

Encrypt



Search

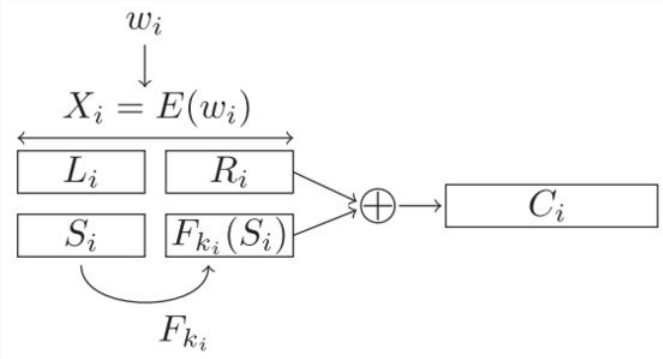
Bosch et al. (2014)



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

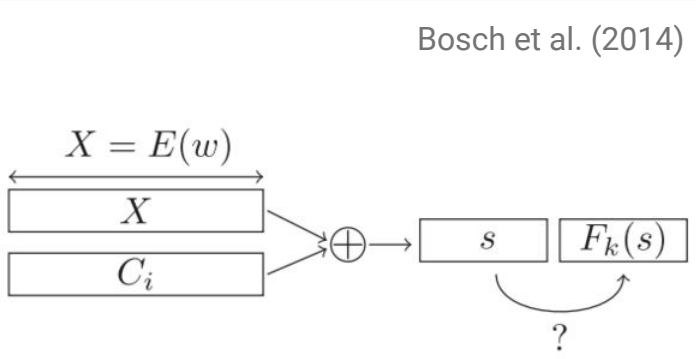
Searchable Symmetric Encryption (SSE)
Song, Wagner, Perrig (SWP)

Encrypt



Search

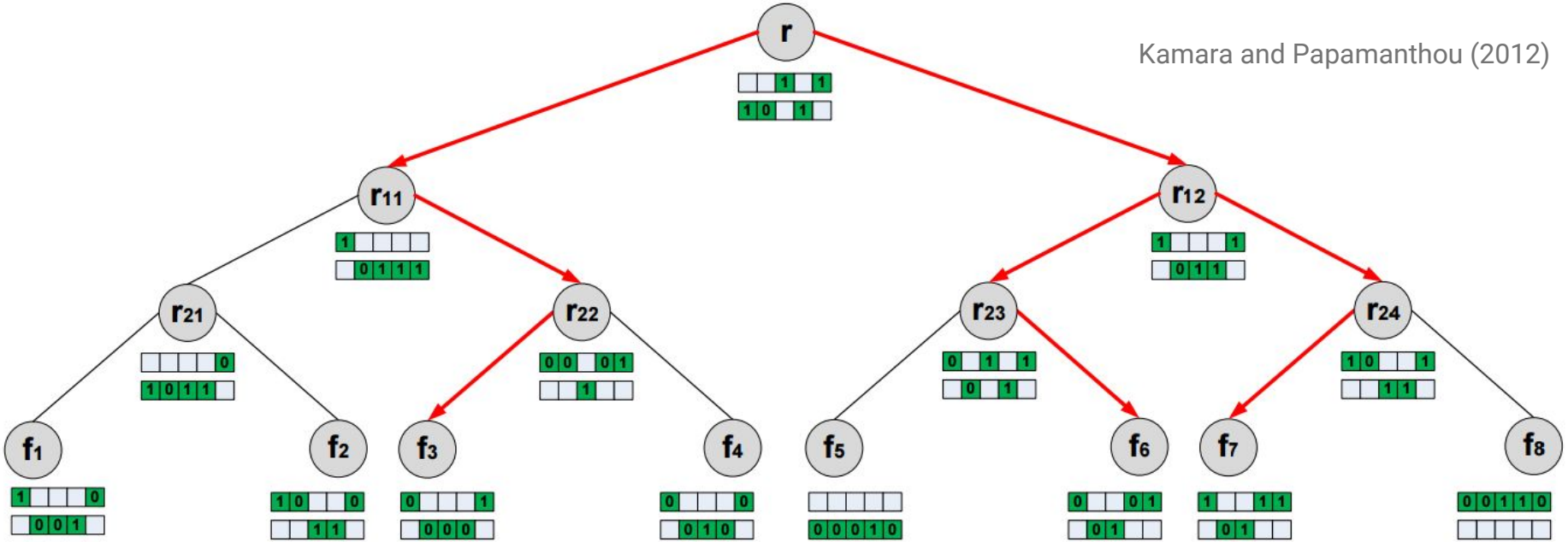
Bosch et al. (2014)



➔ Slow search time: *O(#word * #documents)*, but we can index this

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

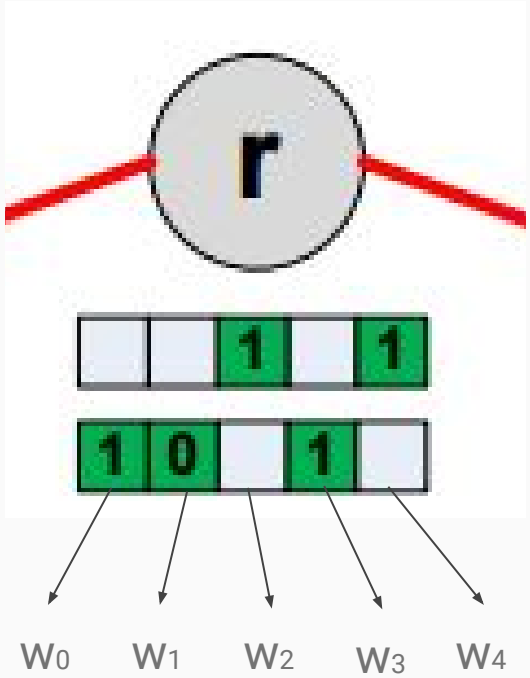
Dynamic SSE (DSSE)
Kamara, Papamanthou (KP)



Kamara and Papamanthou (2012)

- ➔ Uses a data structure called keyword red-black (KRB) tree
- ➔ Leaf nodes contain document ids

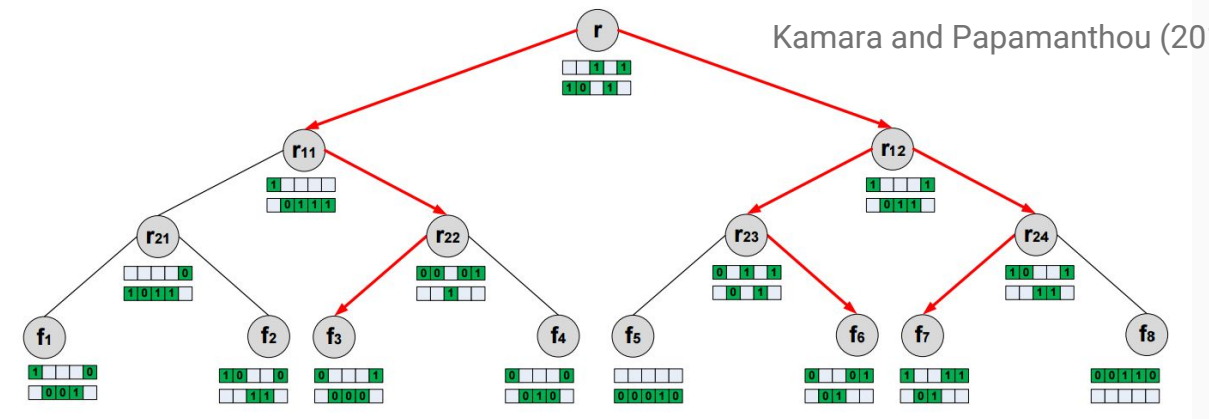
Dynamic SSE (DSSE)
Kamara, Papamanthou (KP)



- The i-th item in the bit array accounts for the i-th word
- If $bit_i = 1$, then there is a path to a file matching the i-th word from this node

- The bit array is split into two arrays
- Randomly assign the correct values 0 and 1 into the one of the two arrays.
- Derive a secret key, sk_i, for each word from some master key
- Encrypt each value in both arrays with sk_i

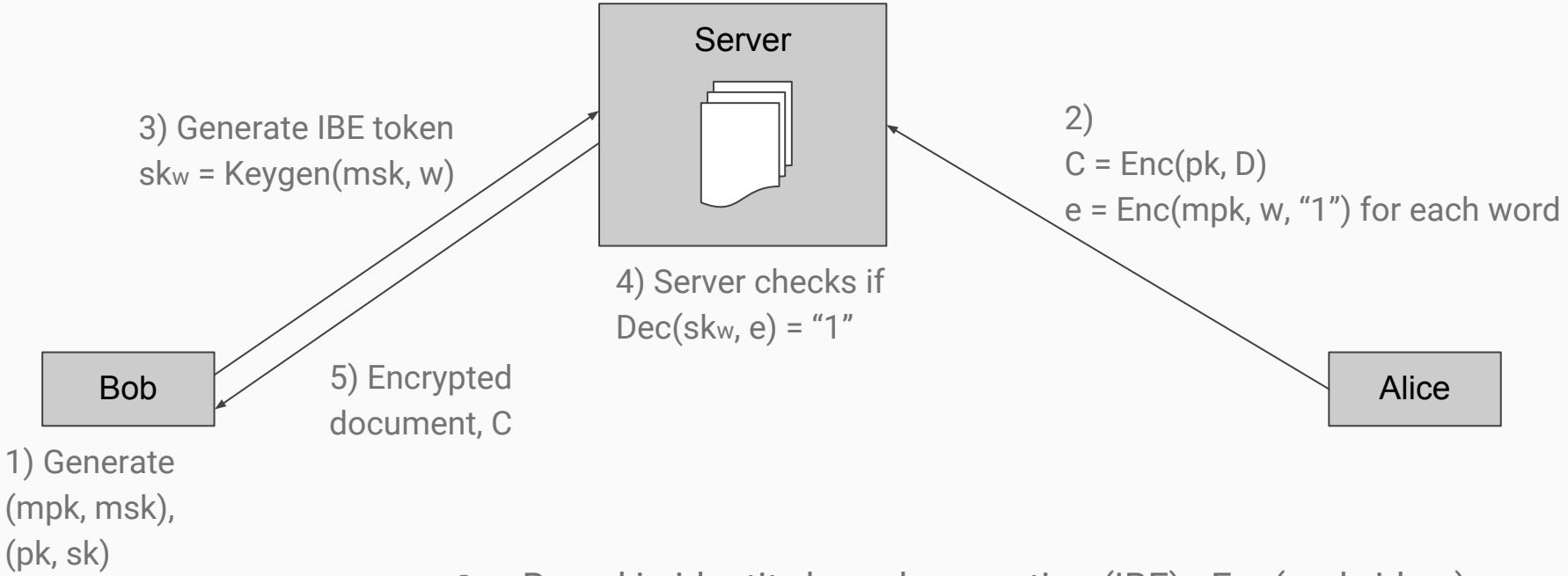
Parallel and Dynamic SSE
Kamara, Papamanthou (KP)



Kamara and Papamanthou (2012)

- ➔ Authors prove that this is CKA2 secure
- ➔ Search time is *O(r log n)*
  - ◆ where r = number of documents containing a keyword and n = number of documents
- ➔ Allows dynamic updating, add, and deleting files
- ➔ Large index size of *O(#documents * #words)*

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



- ➔ Based in identity based encryption (IBE) - Enc(mpk, id, m)
- ➔ Search time is also *O(#documents * #keywords)*, but can be improved with indexing.

Comparisons

scheme	security	search time
SWP (SSE)	CPA	O(mn)
KP (DSSE)	CKA2	O(r log n)
BCO (PEKS)	CKA2	O(mn)

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

Conclusion

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

References

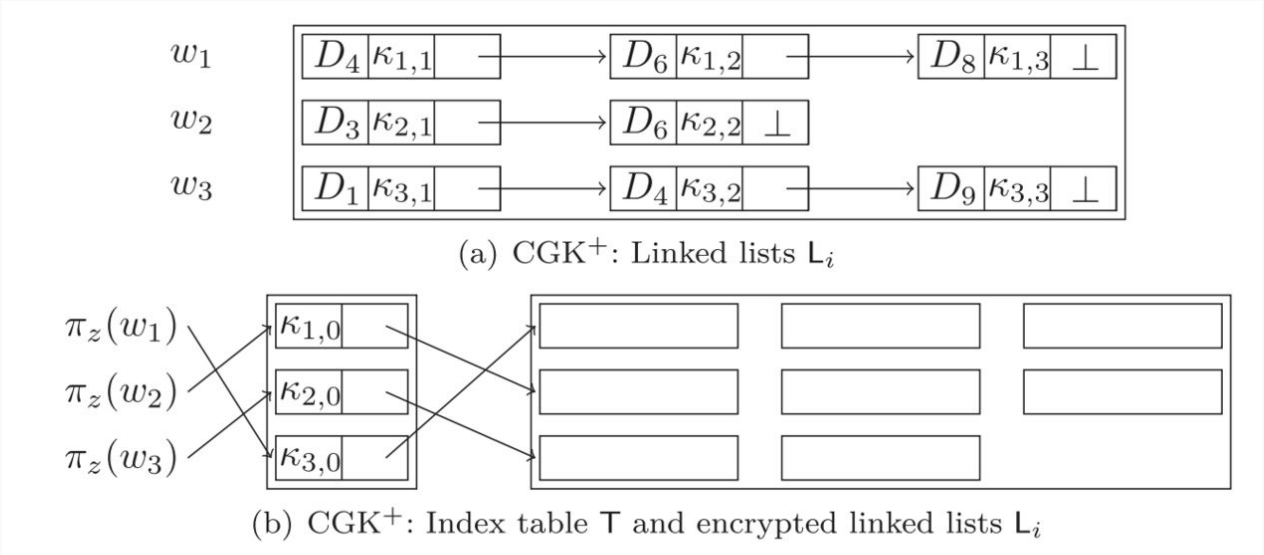
1. D.X. Song, D. Wagner, and A. Perrig. *Practical Techniques for Searches on Encrypted Data*. In 2000 Proceedings of the 2000 IEEE Symposium on Security and Privacy. 2000.

2. Seny Kamara and Charalampos Papamanthou. Parallel and Dynamic Searchable Symmetric Encryption. Financial Cryptography and Data Security. 2013.

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.

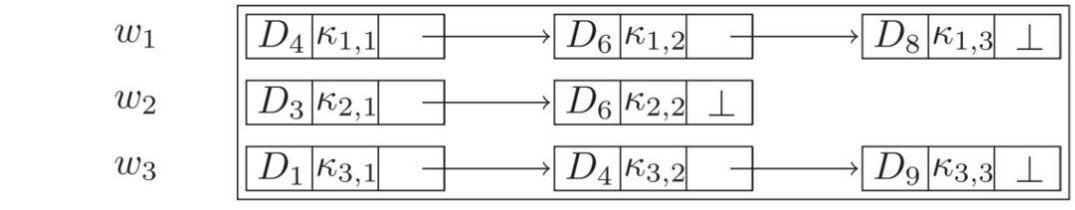
Searchable Symmetric Encryption (SSE)
Curtmola, Garay, Kamara (CGK)



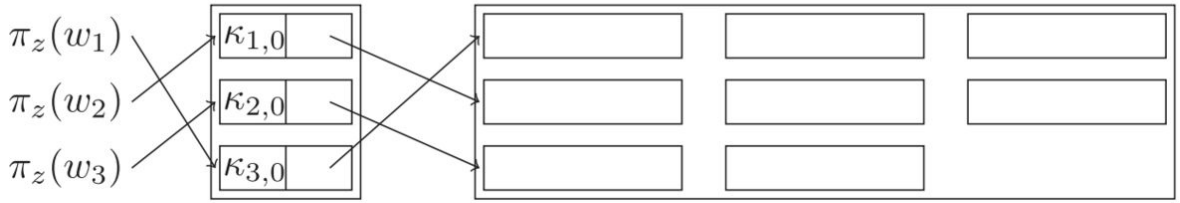
- ➔ Inverted index
- ➔ Consists of a linked list per distinct keyword
- ➔ Each node contains the document id and the key used to encrypt the next node
- ➔ Lookup table that maps the value of a PRF with some key z of the keyword to the head node

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

- ➔ Fast: Sublinear search time O(r)
- ➔ High overhead
- ➔ Difficult to update documents

Bosch et al. (2014)