Property-preserving encryption

Tom Ristenpart
CS 6431
Summary of crypto so far

• ORAM (guest lecture by Elaine)
  – Outsourced memory that hides access patterns, content
• Standard symmetric encryption
• Today:
  – Functionality-preserving encryption
Symmetric encryption

R signifies fresh per-message random bits.

Correctness: \( \text{Dec}(K, \text{Enc}(K,M,R)) = M \) with probability 1 over randomness used
Chosen-plaintext SE security (IND$)

Encryption algorithm denoted Enc

Adversary gets to submit messages to oracle

Ind$ challenge game

Encrypt(M)
C' <- $ Enc(K,M)
If b == 1 then
   C <- C'
If b == 0 then
   C <- $ {0,1}^|C'|
Ret C

b' <- M_1 -----> C_1 -----> M_2 -----> C_2 -----> ...

Adversary outputs guess b' of b
WINS if b' = b

Security goal: Enc(K,M) looks like random bit string to attackers that can obtain encryptions of chosen plaintexts

b is a uniformly sampled bit and K is uniformly sampled key
Both hidden from adversary
Outsourced storage settings

What storage service provider functionalities broken?

• Keyword search (find all files with “Tom” in text)
• Deduplication (if two files same, store only one copy)
• Thumbnails
• ...
Outsourced storage settings

Salesforce stores customer records for companies

Name: Alice
Age: 22
Salary: 100,000
Comments: Works in NYC office
Gender: Female
SSN: 555-31-4325

What storage service provider functionalities broken?

• Field search  (Find all records with Name = Alice)
• Keyword search
• Range queries  (Find all people who make between 90k and 120k)
• Format problems  (Age must be integer between 0 and 130)
• ...
Lots of industry activity

Cloud access security brokers (CASB)
# Functionality-preserving encryption

<table>
<thead>
<tr>
<th>Problem</th>
<th>Crypto primitive</th>
<th>Description</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword search</td>
<td>Searchable symmeric encryption</td>
<td>Perform search over ciphertexts given encrypted search token</td>
<td>[Dawn, Song, Wagner 2000] [Curtmola et al. 2006]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td>Equality search</td>
<td>Deterministic encryption</td>
<td>$X = Y$ implies $\text{DetEnc}(X) = \text{DetEnc}(Y)$</td>
<td>[Rogaway Shrimpton 06]</td>
</tr>
<tr>
<td>Range queries</td>
<td>Order-preserving encryption</td>
<td>$X &gt; Y$ implies $\text{OPEnc}(X) &gt; \text{OPEnc}(Y)$</td>
<td>[Boldyreva et al. 2009]</td>
</tr>
<tr>
<td>Deduplication</td>
<td>Message-locked encryption (Convergent encryption)</td>
<td>Different user’s encryptions of same plaintext give same ciphertext</td>
<td>[Douceur et al. 2002] [Bellare et al. 2013]</td>
</tr>
<tr>
<td>Format restrictions</td>
<td>Format-preserving encryption</td>
<td>Ciphertext has same format as plaintext</td>
<td>[Bellare et al. 2009]</td>
</tr>
</tbody>
</table>
Encryption “onions”

Name: C1 = Enc(K1, DetEnc(K’, “Alice” ) )
Age: C2 = Enc(K2, OPEnc(K’, “22” ) )
Salary: C3 = Enc(K3, OPEnc(K’, “100,000” ) )
Encryption “onions”

- Name: \( C_1 = \text{Enc}(K_1, \text{DetEnc}(K', \text{“Alice”} )) \)
- Age: \( C_2 = \text{Enc}(K_2, \text{OPEnc}(K', \text{“22”} )) \)
- Salary: \( C_3 = \text{Enc}(K_3, \text{OPEnc}(K', \text{“100,000”} )) \)

At this point nothing is leaked to Database:
- Full symmetric encryption security given by \( \text{Enc} \)
- Repeats not visible, ordering not visible
Perform equality search by:
- Send outer encryption key for searched-on-column
- Server decrypts that column
- Use equality preservation of DetEnc to find rows corresponding to Alice

Example:

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alice</td>
<td>22</td>
<td>100,000</td>
</tr>
</tbody>
</table>

K1, DetEnc(K', "Alice")
Perform equality search by:
- Send outer encryption key for searched-on-column
- Server decrypts that column
- Use equality preservation of DetEnc to find rows corresponding to Alice

At this point equality patterns of first column leaked
Perform range search by:

- Send outer encryption key for searched-on-column
- Server decrypts that column
- Use order preservation of OPEnc to find rows in range
Perform range search by:

- Send outer encryption key for searched-on-column
- Server decrypts that column
- Use order preservation of OPEnc to find rows in range

At this point equality patterns of first and second column leaked
Ordering information on second column leaked
CryptDB

• Lots of engineering details
  – Other types of queries, other kinds of onions
  – How to decompose SQL queries into appropriate levels of encrypted data

• Showed that many columns need not be decrypted in some workloads

• Big open question:
  How damaging is the leakage?