Key derivation, deterministic encryption, SIV, wide PRP, tweakable encryption, and format preserving encryption

Pawel Wocjan

Department of Electrical Engineering & Computer Science
University of Central Florida
wocjan@eecs.ucf.edu

October 21st, 2013
Key derivation
Deriving multiple keys from a single key

typical scenario: a single source key (SK) sampled from:

- hardware random number generator
- key exchange protocol

we need many keys to secure a session:

- unidirectional keys
- multiple keys for nonce-based CBC

goal: generate multiple keys from a single source key

\[
\text{SZ} \xrightarrow{\text{KDF}} k_1, k_2, k_3, \ldots
\]
Source key SK is uniform

$F$: a PRF with key space $\mathcal{K}$ and output in $\{0, 1\}^n$

suppose source key SZ is uniform in $\mathcal{K}$

define key derivation function KDF as

$$KDF(SK, CTX, L) = F(SK, (CTX||0)) \| F(SK, (CTX||1)) \| \cdots F(SK, (CTX||L - 1))$$

$CTX$: a string that uniquely identifies the application
Source key SZ is not uniform

recall that PRFs are pseudo random only if the key is uniform in $\mathcal{K}$

- SK not uniform $\Rightarrow$ PRF output does not look random

source key often not uniform:

- key exchange protocol: key uniform in a subset of $\mathcal{K}$
- hardware RNG: may produce biased output
Extract-then-Expand paradigm

1. step: extract pseudo-random key $k$ from source key $SZ$

2. step: expand $k$ by using it as a PRF key as before
HKDF: KDF from HMAC

implements the extract-then-expand paradigm:

- extract: use \( k = \text{HMAC}(\text{salt}, SZ) \)

  salt: a fixed non-secret string chosen at random

- expand using HMAC as a PRF with key \( k \)
Password-based KDF (PBKDF)

deriving keys from passwords

- do not use HKDF: passwords have insufficient entropy
- derived keys would be vulnerable to dictionary attacks

PBKDF defenses: salt and a slow hash function

standard approach: PKCS#5

\[ H^{(c)}(pwd||salt) : \text{iterate hash function } c \text{ times} \]
Deterministic encryption
The need for deterministic encryption

- deterministic encryption enables server to make lookup later
- for instance, serves ask DB to retrieve record \( E(k_1, \text{Alice}) \) and receives \( E(k_2, \text{stuff}) \)
Problem: deterministic encryption cannot be CPA secure

the attacker can tell when two ciphertexts encrypt the same message
⇒ leaks information
⇒ leads to significant attacks when message space $\mathcal{M}$ is small
Problem: deterministic encryption cannot be CPA secure

1. adversary submits $m_0, m_0 \in \mathcal{M}$ and obtains $c_0 = E(k, m_0)$ from challenger

2. adversary submits $m_0, m_1 \in \mathcal{M}$ and obtains $c = E(k, m_b)$ from challenger

3. adversary outputs 0 if $c = c_0$ and 1 otherwise
A solution: the case of unique messages

suppose encryptor never encrypts same message twice: the pair 
\((k, m)\) never repeats

this happens when encryptor:

- chooses messages at random from a large message space (e.g. keys)
- message structure ensure uniqueness (eq. unique user id)
Deterministic CPA security

let \((E, D)\) be a cipher defined over \((\mathcal{K}, \mathcal{M}, \mathcal{C})\)

for \(b = 0, 1\) define \(W_b\) as

[challenger]

receives \(b\)

\(k \leftarrow \mathcal{K}\)

for \(i = 0, \ldots, q - 1\)

\(m_{0,i}, m_{1,i} \in \mathcal{M} : |m_{0,i}| = |m_{1,i}|\)

\(c_i = E(k, m_{b,i})\)

adversary

outputs \(b'\)

\(m_{0,0}, \ldots, m_{0,q-1}\) are distinct and \(m_{1,0}, \ldots, m_{1,q-1}\) are distinct

\(E\) is semantically secure under deterministic CPA \(\iff\) \(\forall\) efficient \(A\)

\[\text{Adv}_{d\text{CPA}}(A, E) = |\Pr(W_0 = 1) - \Pr(W_1 = 1)|\] is negligible
Is CBC with fixed IV is not deterministic CPA secure?
Is counter mode with fixed IV deterministic CPA secure?
Deterministic encryption:
SIV and wide PRP
Deterministic encryption

needed for maintaining an encrypted database index

- lookup records by encrypted index
deterministic CPA security

- no message is ever encrypted with the same key:
  the pair \((k, m)\) is always unique

we defined deterministic CPA security game
Construction 1: Synthetic IV

- Let $(E, D)$ be a CPA-secure cipher with randomness $k \in \mathcal{K}_1$, $m \in \mathcal{M}$, $r \in \mathcal{R}$, and $c = E(k, m; r) \in \mathcal{C}$

- Let $F : \mathcal{K}_2 \times \mathcal{M} \rightarrow \mathcal{R}$ be a secure PRF

- Define $E_{\text{det}} : (\mathcal{K}_1 \times \mathcal{K}_2) \times \mathcal{M} \rightarrow \mathcal{M}$ by

\[
E_{\text{det}}((k_1, k_2), m) = E(k_2, m; F(k_1, m))
\]

- Theorem: $E_{\text{det}}$ is semantically secure under deterministic CPA
  proof sketch: distinct messages $\Rightarrow$ all $r$'s generates from messages are indistinguishable from random

- SIV is well suited for messages longer than one AES block
Ciphertext integrity

- goal: deterministic CPA security and ciphertext integrity $\Rightarrow$
  
  **DAE**: deterministic authenticated encryption

- consider the special case SIV-CTR

$$\begin{align*}
F_{\text{CTR}}(k_2, IV) \parallel F_{\text{CTR}}(k_2, IV + 1) \parallel \cdots \parallel F_{\text{CTR}}(k_2, IV + (L - 1))
\end{align*}$$
Deterministic authenticated encryption

Theorem: if $F$ is secure PRF and CTR from $F_{CTR}$ is CPA-secure $\Rightarrow$ SIV-CTR based on $F$ and $F_{CTR}$ provides DEA
Construction 2: simply use a PRP

- let \((E, D)\) be a secure PRP, \(E : \mathcal{K} \times \mathcal{X} \rightarrow \mathcal{X}\)

- Theorem: \((E, D)\) is semantically secure under deterministic CPA

  proof sketch: let \(f : \mathcal{X} \rightarrow \mathcal{X}\) be a truly random invertible function

  - \(b = 0 \Rightarrow \) adversary sees \(f(m_0, 0), \ldots, f(m_0, q-1)\)
  - \(b = 1 \Rightarrow \) adversary sees \(f(m_1, 0), \ldots, f(m_1, q-1)\)

- AES yields deterministic CPA secure encryption for only 16 byte messages

we need to construct PRP for longer messages
EME: constructing a wide block PRP
goal: deterministic CPA security and ciphertext integrity

message $\parallel 00\ldots0$ $\xrightarrow{E(k,\cdot)}$ ciphertext

message $\parallel *\ldots*$ $\xrightarrow{D(k,\cdot)}$ ciphertext

if $\not= 00\cdots0$ output $\bot$

Theorem: let $(E, D)$ be a secure PRP with

$E : \mathcal{K} \times (\mathcal{M} \times \{0, 1\}^n) \to (\mathcal{M} \times \{0, 1\}^n)$

$1/2^n$ negligible $\Rightarrow$ PRP-based encryption provides DEA
Tweakable encryption
Format preserving encryption