CS 6160 Cryptology Lecture 10: CPA-security & Modes of Operation

Maria Francis

October 06, 2020

CPA-Secure Encryption from PRFs

- We continue our discussion of CPA-secure schemes.
- Reading 3.5.2 of Katz and Lindell textbook
- We saw Pseudorandom Functions and the natural question to ask is : can I have CPA-secure encryption schemes using PRFs?
- We just need to construct a scheme for fixed-length, we already know that implies for arbitrary length messages as well!
- One way: $Enc_k(m) = F_k(m)$, but it is deterministic and cannot be CPA-secure.
- Instead we give a random value r as input to PRF and XOR output with plaintext.

CPA-Secure Encryption Algorithm

- Let F be a PRF and message length is n.
- *Gen*:1^n is its input, it chooses $k \in \{0,1\}^n$ and outputs it.
- *Enc*: on input a key $k \in \{0,1\}^n$ and $m \in \{0,1\}^n$, choose uniform $r \in \{0,1\}^n$ and output,

$$c := \langle r, F_k(r) \oplus m \rangle$$

- *Dec*: on input $k \in \{0,1\}^n$ and $c = \langle r, s \rangle$, it outputs

$$m := F_k(r) \oplus s$$

CPA-Security Proof

- Let $\overline{\Pi} = (\overline{Gen}, \overline{Enc}, \overline{Dec})$ be the identical encryption scheme to Π except that a truly random function $f \in Func_n$ is used in place of F_k.
- An $\mathcal{A}$ and $q(n)$: upper bound on the number of queries that $\mathcal{A}$ makes to the Enc-oracle.
- Claim1: there exist a $\mathrm{negl}(n)$ s.t.:

$$|Pr[PrivK^{cpa}_{\mathcal{A},\Pi}(n) = 1] - Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1]| \leq \mathrm{negl}(n).$$

- Proof by Reduction:
  - We use $\mathcal{A}$ to construct a distinguisher D – that has an oracle access to some function $\mathcal{O}$ – for the PRF F.
  - D runs the $PrivK^{cpa}$ experiment and see if $\mathcal{A}$ succeeds. If it does then D guesses that $\mathcal{O}$ is a PRF, else it is f.

Distinguisher D

- D is given input 1^n and access to oracle $\mathcal{O} : \{0,1\}^n \to \{0,1\}^n$.

1. Run $\mathcal{A}(1^n)$. $\mathcal{A}$ queries the Enc-oracle on a message $m \in \{0,1\}^n$, answer this query:
   (a) Choose a uniform $r \in \{0,1\}^n$.
   (b) Query $\mathcal{O}(r)$ and obtain response y.
   (c) Return $\langle r, y \oplus m \rangle$ to $\mathcal{A}$.
2. $\mathcal{A}$ outputs messages $m_0, m_1 \in \{0,1\}^n$, choose a unform bit $b \in \{0,1\}$ and then:
   (a) Choose uniform $r \in \{0,1\}^n$.
   (b) Query $\mathcal{O}(r)$ and obtain response y.
   (c) Return $\langle r, y \oplus m_b \rangle$ to $\mathcal{A}$
3. Continue answering Enc-oracle queries of $\mathcal{A}$ until $\mathcal{A}$ outputs a bit $b^{'}$. Output 1 if $b^{'} = b$, else 0.

Distinguisher D

- D runs in polynomial times since $\mathcal{A}$ does.
- If D's oracle is a PRF then the view of $\mathcal{A}$ when run as a subroutine by D is distributed identically to the view of $\mathcal{A}$ in experiment $PrivK^{cpa}_{\mathcal{A},\Pi}(n)$.
- Why? All the steps that D does is the same: choosing r, computing $y := F_k(r)$ and $c = \langle r, y \oplus m \rangle$.
- If D's oracle is a random function then the view of $\mathcal{A}$ now is the same as in the experiment $PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n)$.
- By the definition of F being a PRF, we have the claim (Claim 1).

CPA-security proof contd.

- Claim 2:

$$Pr[PrivK^{cpa}_{\mathcal{A},\Pi}(n) = 1] \leq \frac{1}{2} + \frac{q(n)}{2^n}.$$

- Every time m is encrypted in $PrivK^{cpa}_{\mathcal{A},\Pi}(n)$ (either by Enc-oracle or challenge ciphertext), a uniform r is chosen and ciphertext is set $\langle r, f(r) \oplus m \rangle$.
- Let r^* be used for challenge ciphertext. There are two cases:
  1. r^* is never used when answering any of $\mathcal{A}$'s Enc-oracle queries:
     - $\mathcal{A}$ learns nothing about $f(r^*)$ by interacting with Enc-oracle.
     - For $\mathcal{A}$, $f(r^*) \oplus m_b$ is uniformly distributed and independent of the experiment so probability that $b = b'$ is $1/2$.

CPA-security proof contd.

2. r^* came up at least once in $\mathcal{A}$'s Enc-oracle queries.
   - $\mathcal{A}$ gets $\langle r^*, s\rangle$ as response for m, $\Rightarrow f(r^*) = s \oplus m$.
   - Probability of that happening: $q(n)/2^n$, $r^* \in \{0,1\}^n$.

Let *Repeat* be the event corresponding to 2.

$$
\begin{aligned}
&Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1] \\
&= Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1 \cap Repeat] + \\
&Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1 \cap \overline{Repeat}] \\
&\leq Pr[Repeat] + Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1 | \overline{Repeat}] \\
&\leq q(n)/2^n + 1/2.
\end{aligned}
$$

Block-Cipher Modes of Operation

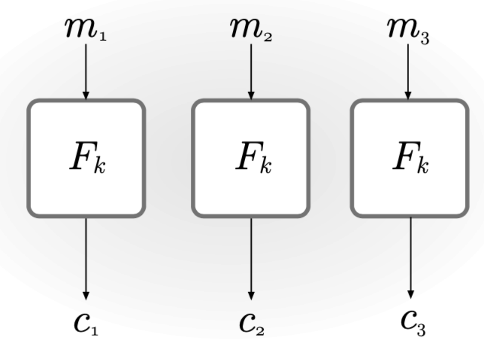
- Modes of operation provide a way to encrypt arbitrary-length messages using shorter ciphertexts.
- Reading : Section 3.6 of Katz & Lindell
- Reading exercise: Stream-ciphers modes of operation
- Let F be a block cipher with blocklength n
- For ease we assume that all messages m have a multiple of n length. Or append with 1 followed by 0s.
- Let $m = m_1, m_2, \ldots, m_\ell,\ m_i \in \{0,1\}^n$ be the plaintext.

Electronic Code Book (ECB) mode

- Naive mode of operation, only historical significance.
- Apply directly the block cipher on each plaintext block. That is

$$c := \langle F_k(m_1), F_k(m_2), \ldots, F_k(m_l) \rangle$$

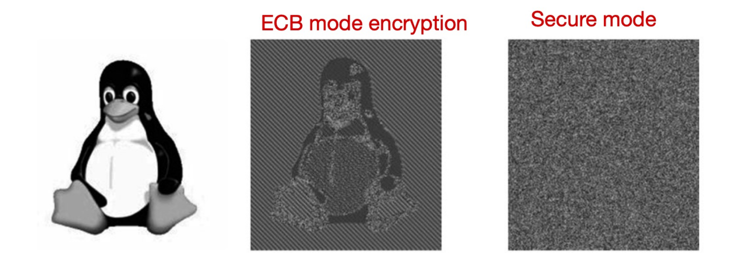
- Decryption is done in the obvious way.



ECB mode - Problems

- It is deterministic and so it is not CPA-secure.
- It does not even have indistinguishable encryptions in the presence of an eavesdropper.
- Main issue: If a block is repeated in the plaintext then the block is repeated in ciphertext.
- Like in previous lecture, we can easily create a indistinguishability experiment that an $\mathcal{A}$ will succeed with certainity: Two plaintexts, one with two identical blocks and other with distinct plaintext blocks.
- Note: not a theoretical problem.
- Encrypting an image where a small group of pixels are now a plaintext block.
- Encrypting with ECB reveals a lot of information when there are repeating patterns.

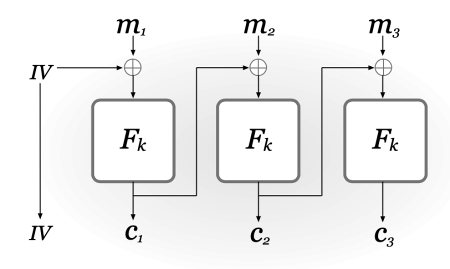
ECB mode - Image encryption issues



Cipher Block Chaining (CBC) mode

- A uniform initialization vector (IV) of length n is first chosen.
- Apply the block cipher to the XOR of the current plaintext block and previous ciphertext block.
- I.e., set $c_0 := IV$ and for $i = 1, \ldots, \ell$, set $c_i := F_k(c_{i-1} \oplus m_i)$
- Decryption of ciphertext $c_0, \ldots, c_\ell$: compute $m_i := F_k^{-1}(c_i) \oplus c_{i-1}$, for $i = 1, \ldots, \ell$.
- Note that IV has to be included in the ciphertext for correct decryption.
- This is called stateful encryption. Its operation depends on a quantity called the state (previous ciphertext) which is prespecified.

CBC mode

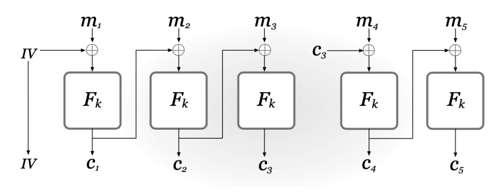


CBC mode

- Encryption is probabilistic.
- It can be proven if F is a pseudorandom permutation then CBC-mode encryption is CPA-secure.
- Issue: Encryption has to be carried out sequentially. Why? c_{i-1} is needed in order to encrypt m_i.
- What if we use a distinct IV everytime instead of a random IV?, i.e. first use $IV = 1$ and then increment.
- Practice Question (3.20): Such a variant of CBC-mode is not secure.

Chained CBC mode

- Last block of ciphertext is used as *IV* for encrypting the next message.
- This reduces bandwidth as *IV* need be sent each time.



- This shows for m_1, m_2, m_3 being encrypted by a random *IV* and then for m_4, m_5 c_3 is the *IV*.
- It is used in SSL 3.0 and TLS 1.0.

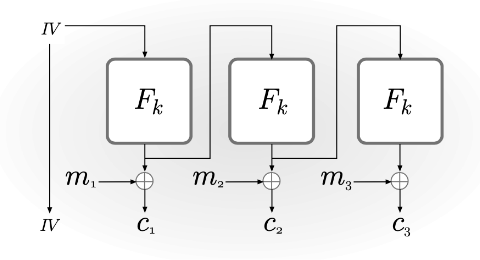
Chained CBC mode - CPA attack

- Basis of this attack is that $\mathcal{A}$ knows the IV that will be used for the second encrypted message.
- Assume that $\mathcal{A}$ knows that $m_1 \in \{{m_1}^0, {m_1}^1\}$ and eavesdrops $IV,\ c_1, c_2, c_3$.
- $\mathcal{A}$ requests for an encryption of a second message m_4, m_5 with $m_4 = IV \oplus m_1^0 \oplus c_3$.
- $\mathcal{A}$ can verify that $m_1 = {m_1}^0$ iff $c_4 = c_1$.
- Seemingly small modifications make the scheme vulnerable to an attack!

Output Feedback Mode (OFB)

- First a uniform $IV \in \{0,1\}^n$ is chosen.
- We generate a pseudorandom stream from IV:
  - Define $y_0 := IV$ and set the ith block y_i as $y_i := F_k(y_{i-1})$.
- Then you XOR each block of plaintext with the appropriate block of the stream: $c_i := y_i \oplus m_i$.
- IV has to be included in the ciphertext for decryption.
- Here unlike CBC F need not be invertible.

Output Feedback Mode (OFB)

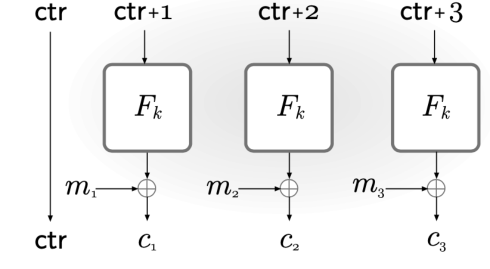


Output Feedback Mode (OFB)

- Plaintext length need not be a multiple of the block length.
- The generated stream can be truncated to exactly the plaintext length.
- Another advantage: final value y_ℓ used to encrypt some message can be used as IV for encrypting the next message and is still secure.
- This is called the stateful variant.
- OFB mode is CPA-secure if F is a PRF.
- Enc and Dec have to be done sequentially, but the bulk of the computation, pseudorandom stream, can be done independent of the actual message, so preprocessing is possible.

Counter (CTR) mode

- A uniform value $\mathrm{ctr} \in \{0,1\}^n$ is first chosen.
- A pseudorandom stream is generated by computing $y_i := F_k(\mathrm{ctr} + i)$, addition is done modulo 2^n.
- The ith ciphertext block $c_i := y_i \oplus m_i$, IV is sent as part of ciphertext.



CTR mode

- CTR Enc and Dec can fully be parallelized, since all blocks are independent of each other.
- Unlike OFB, it is possible to decrypt the ith block of the ciphertext using only one evaluation of F.
- It is an attractive choice and its security is easy to analyze too!

Theorem
If F is PRF, then CTR mode is CPA-secure.

CTR mode - CPA security

- Let $\Pi = (Gen, Enc, Dec)$ be the CTR mode encryption scheme with F_k as PRF.
- Let $\overline{\Pi} = (\overline{Gen}, \overline{Enc}, \overline{Dec})$ be the identical encryption scheme to Π except that a truly random function $f \in Func_n$ is used in place of F_k.
- Let $\mathcal{A}$ be a PPT adversary and $q(n)$: the polynomial upper bound on Enc-oracle queries made by $\mathcal{A}(1^n)$, maximum no of blocks in any such query and max number of blocks in m_0 and m_1.
- We use our previous result about PRFs to claim that there exist $\mathrm{negl}(n)$ s.t.:

$$|Pr[PrivK^{cpa}_{\mathcal{A},\Pi}(n) = 1] - Pr[PrivK^{cpa}_{\mathcal{A},\overline{\Pi}}(n) = 1]| \leq \mathrm{negl}(n).$$ Prove!

CTR mode - CPA security

- Slightly non-trivial part, showing:

$$Pr[PrivK^{cpa}_{\mathcal{A},\Pi}(n) = 1] < \frac{1}{2} + \frac{2q(n)^2}{2^n}.$$

- Combining previous equation:

$$Pr[PrivK^{cpa}_{\mathcal{A},\Pi}(n) = 1] < \frac{1}{2} + \frac{2q(n)^2}{2^n} + \mathrm{negl}(n).$$

- Since q is polynomial, $\frac{2q(n)^2}{2^n}$ is negligible and we are done!

CPA Security Proof

- What we just saw is a standard technique for CPA-security.
- The first step of such proofs is to consider a hypothetical version of the construction in which the PRF is replaced with a $f \in Func_n$.
- Then using a proof by reduction we argue that this modification does not significantly affect the attacker's success probability.
- We only now need to analyze the scheme that uses a completely random function. This would rely typically only on probabilistic analysis and not on any computational assumptions.

Modes of Operation and Message Tampering

- Here we only look at encryption or data confidentiality.
- Message tampering looks at message integrity or authentication which is studied separately in Katz & Lindell textbook and we will follow that.
- None of the modes we discussed achieve message integrity.
- Practice Questions 3.21 and 3.22 look at errors that might come in during transmission and not because of adversarial interference.

Block Length and Concrete Security

- CBC, OFB, CTR use random IV to randomize Enc and ensure that the block cipher works on fresh inputs.
- This is key for CPA-security.
- So now block length has an impact:
  - In a CTR mode using a block cipher F with ℓ block length.
  - IV is then a uniform ℓ-bit string and we expect an IV to repeat after encrypting $2^{\ell/2}$ messages. Read about the Birthday Problem.
  - If ℓ is too short, even if F is a PRF the concrete security is too weak for practical applications.
  - If $\ell = 64$ as in the case of DES, then after 2^{32} encryptions $\approx 34GB$ of plaintext a repeated IV can happen.
  - It is not a lot of data!

IV misuse

- What if it is not a random IV?
- For OFB and CTR it is much worse than CBC.
- Why? If an IV repeats then in OFB and CTR $\mathcal{A}$ can XOR the two ciphertexts and get info about entire contents of both the encrypted messages.
- In CBC mode, it is likely that after only a few blocks the inputs to the block cipher will diverge and $\mathcal{A}$ will be unable to learn any information beyond the first few message blocks.
- Stateful/chained mode where dependency on IV is reduced is a workaround.