Block ciphers (2)

CS 5830

Blockciphers

Family of permutations, one permutation for each key

$$E : \{0,1\}^k \times \{0,1\}^n \text{-->} \{0,1\}^n$$

Use notation $E(K,X) = E_K(X) = Y$

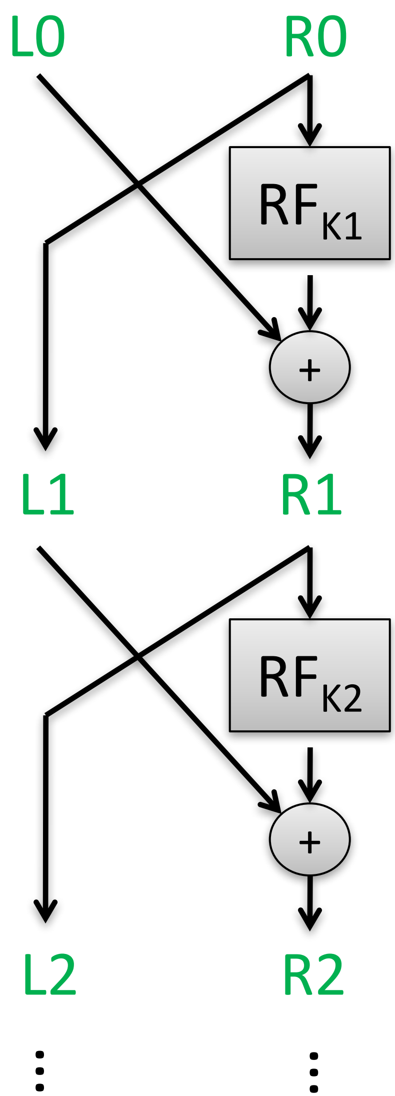
Define inverse $D(K,Y) = D_K(Y) = X$ such that $D_K(E_K(X)) = X$

E,D must be efficiently computable

Key generation: pick K uniformly at random from $\{0,1\}^k$

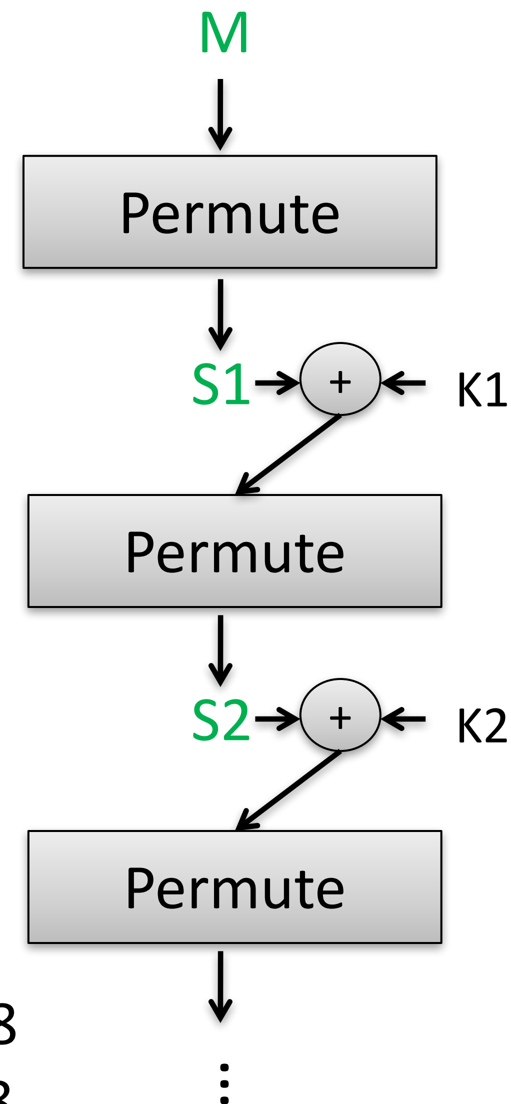
Nowadays $k \geq 128$

DES and AES blockciphers



n = 64
k = 56

n = 128
k = 128



PRF definition

Adversary (distinguisher) can't tell between E_K and random function

Let Func(n,n) be set of all functions from n bits to n bits

$$\epsilon = \left| \Pr[K \leftarrow \{0,1\}^k \ : \ \mathcal{A}^{E_K(\cdot)} = 1] - \Pr[\rho \leftarrow Func(n,n) \ : \ \mathcal{A}^{\rho(\cdot)} = 1] \right|$$

Select a secret key

Select a random function

- Insecure if we can find adversary $\mathcal{A}$ such that ϵ is close to 1
- Secure if we can prove that no (computationally efficient) adversary can achieve advantage far from 0
- Adversary is computationally bound in run time, and consequently can only make a limited number of queries to its oracle

Permutations as PRFs

Adversary (distinguisher) can't tell between E_K and random function

Let Func(n,n) be set of all functions from n bits to n bits

$$\epsilon = \left| \Pr[K \leftarrow \{0,1\}^k \; : \; \mathcal{A}^{E_K(\cdot)} = 1] - \Pr[\rho \leftarrow Func(n,n) \; : \; \mathcal{A}^{\rho(\cdot)} = 1] \right|$$

Select a secret key

Select a random function

- Recall that we require E_K to be a *permutation*
- Can we give an adversary $\mathcal{A}$ such that ϵ is close to 1 for any E_K ?

Pseudorandom permutations (PRPs)

Adversary (distinguisher) can't tell between E_K and random permutation

Let Perm(n,n) be set of all permutations from n bits to n bits

$$\epsilon = \left| \Pr[K \leftarrow \{0,1\}^k \ : \ \mathcal{A}^{E_K(\cdot)} = 1] - \Pr[\pi \leftarrow Perm(n,n) \ : \ \mathcal{A}^{\pi(\cdot)} = 1] \right|$$

Select a secret key

Select a random permutation

- Insecure if we can find adversary $\mathcal{A}$ such that ϵ is close to 1
- Secure if we can prove that no (computationally efficient) adversary can achieve advantage far from 0
- Adversary is computationally bound in run time, and consequently can only make a limited number of queries to its oracle

PRP/PRF Switching Lemma

If n is large enough, then not much difference between PRP and PRF. More formally:

$$\left|\Pr[\rho \leftarrow Func(n,n)\ :\ \mathcal{A}^{\rho(\cdot)} = 1] - \Pr[\pi \leftarrow Perm(n,n)\ :\ \mathcal{A}^{\pi(\cdot)} = 1]\right| \leq \frac{q^2}{2^n}$$

where q is number of oracle queries $\mathcal{A}$ makes

n = 4	pretty good attack even with q = 2
n = 64	q must get close to 2^{32} to distinguish
n = 128	q must get close to 2^{64} to distinguish

Encryption from good PRF/PRP

Recall our multi-message encryption:

<u>$\text{Enc}_K(m)$:</u>
$r \leftarrow U_n$
Return ($r, m \oplus E_K(r)$)

This is provably multi-message secure
if E is secure PRF (or, by switching lemma, PRP)

Instantiating PRF with AES

Recall our multi-message encryption:

$\underline{\mathrm{Enc}_K(m):}$
$r \leftarrow U_n$
Return ($r, m \oplus \mathrm{AES}_K(r)$)

This is provably multi-message secure
if AES is secure PRF (or, by switching lemma, PRP)

We will make this assumption, and trust that no cryptanalysts can find better attacks

Two encryption applications

We'll look closely at two encryption applications:

- **Length-preserving encryption**
  - Useful for cases where ciphertexts must be same length as plaintexts.
  - Should only be used when absolutely needed
- **Length-extending encryption (used for TLS)**
  - Insecure variants: CTR mode, ECB mode, CBC mode
  - We'll build secure ones in a few lectures

Example: Credit card number encryption

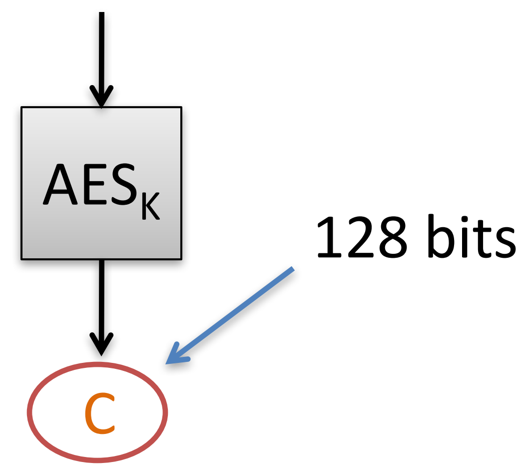
Jane Doe	1343-1321-1231-2310
Thomas Ristenpart	9541-3156-1320-2139
John Jones	5616-2341-2341-1210
Eve Judas	2321-4232-1340-1410

Database schemas and software require <= 16 decimal digits and valid Luhn checksum

$AES_K : \{0,1\}^{128} \longrightarrow \{0,1\}^{128}$

Ciphertexts are too big for replacing plaintext within database!

M = 2321-4232-1345-1415



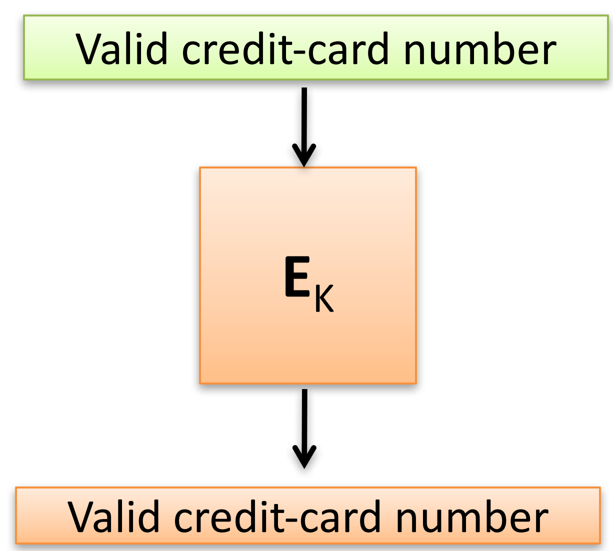
Example: Credit card number encryption

Jane Doe	
Thomas Ristenpart	
John Jones	
Eve Judas	

Database schemas and software require <= 16 decimal digits and valid Luhn checksum

Encryption tool whose ciphertexts are also credit-card numbers

$$E_K : [0..9]^{16} \longrightarrow [0..9]^{16}$$



Format-preserving encryption (FPE)

$$M \longrightarrow \boxed{E_K} \longrightarrow C$$

Enjoys format N   Enjoys format N

Disk sectors / payment card numbers just two examples
Some others:

1) Valid addresses for a certain country

2) 4096-byte disk sectors

3) Assigned Social Security Numbers (9 digits, without leading 8 or 9)

4) Composition of (1) and (3)

How to build FPE on 48 bits?

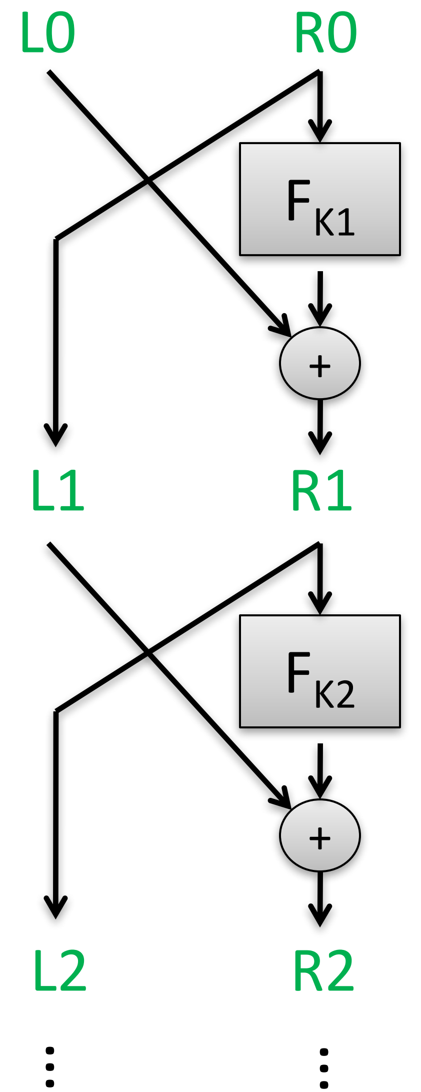
Special case of FFX encryption

Input M = 48 bits
L0 = 24 bits
R0 = 24 bits

$F_{K1}(R) = AES(K, 1 \ || \ R)$
$F_{K2}(R) = AES(K, 2 \ || \ R)$
...

Take XOR mod 2^{24}

Use 10 rounds



Balanced Feistel security in theory

- Luby & Rackoff showed that if round functions are PRFs and n is relatively large, then
  - 3 rounds suffice to prove that Feistel is a PRP
  - Proofs hold up to $q \approx 2^{n/4}$
- For FPE n is often *not very large*:
  - FFX designers suggested 10 rounds as heuristic
  - Recent “certificational” weaknesses against 10 rounds [Bellare, Hoang, Tessaro 2016]

FPE now widely used in practice

Security problems with length-preserving encryption?

$M \longrightarrow \boxed{E_K} \longrightarrow C$

Looks like a random string

But determinism has problems:

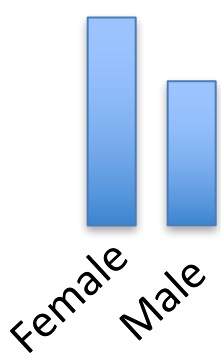
	Plaintext	Ciphertext
Jane Doe	1343-1321-1231-2310	1049-9310-3210-4732
Thomas Ristenpart	9541-3156-1320-2139	7180-4315-4839-0142
John Jones	2321-4232-1340-1410	5731-8943-1483-9015
Eve Judas	1343-1321-1231-2310	1049-9310-3210-4732

Simple frequency analysis attacks

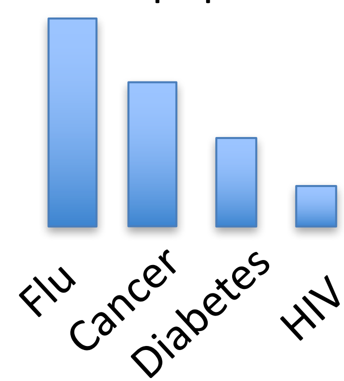
Say adversary steals a medical database with a column encrypted with FPE

Patient #	Sex	Disease type
0	11110010	101010101001000
1	10101100	111110101000101
2	10101100	111110101000101
3	10101100	001111100011111

Know sex is only Male or Female
More women go to hospital then men



Know 4 types of diseases and their distribution population



Simple frequency analysis attacks

Say adversary steals a medical database with a column encrypted with FPE

Patient #	Sex	Disease type
0	11110010	101010101001000
1	10101100	111110101000101
2	10101100	111110101000101
3	10101100	001111100011111

There are some mitigations for attacks, but in general one should use FPE ***only as a last resort***!

Length-extending encryption security

- Not a bit of information about plaintext leaked
  - Equality of plaintexts hidden
  - Even in case of active attacks (we'll get to this)
    - Padding oracles we will see later
- Eventually: authenticity of messages as well
  - Decryption should reject modified ciphertexts