

Stream ciphers

The One Time Pad PRG

Symmetric Ciphers: definition

Def: a cipher defined over $(\mathcal{X}, \mathcal{M}, \mathcal{C})$ is a pair of "efficient" algs (E, D) where $E: \mathcal{X} \times \mathcal{M} \to \mathcal{C}$ $D: \mathcal{X} \times \mathcal{C} \to \mathcal{M}$ S.L. $\forall m \in \mathcal{M}, \kappa \in \mathcal{X}: D(\kappa, E(\kappa, m)) = M$

• E is often randomized. D is always deterministic.

The One Time Pad

(Vernam 1917)

First example of a "secure" cipher

$$\mathcal{M} = G = \{0,1\}^h$$

$$\mathcal{A} = \{0,1\}^h$$

key = (random bit string as long the message)

The One Time Pad

(Vernam 1917)

$$C := E(K, m) = K \oplus M$$

$$D(K, c) = K \oplus C$$

You are given a message (m) and its OTP encryption (c).

Can you compute the OTP key from m and c?

Yes, the key is
$$k = m \oplus c$$
.

The One Time Pad

(Vernam 1917)

```
Very fast enc/dec!!
... but long keys (as long as plaintext)
```

Is the OTP secure? What is a secure cipher?

What is a secure cipher?

Attacker's abilities: CT only attack (for now)

Possible security requirements:

attempt #1: attacker cannot recover secret key

attempt #2: attacker cannot recover all of plaintext

Shannon's idea:

CT should reveal no "info" about PT

Information Theoretic Security

(Shannon 1949)

Def: A cipher
$$(E, D)$$
 over (K, M, C) has perfect secrecy if

 $\forall m_0, m_1 \in M \quad (|vu(m_0)| = |vu(m_1)) \quad \text{and} \quad \forall c \in C$
 $|Pr[E(K, m_0)| = c] = |Pr[E(K, m_1)| = c]$

where $|K|$ is various in $|\mathcal{J}_{K}| = |K| = |K|$

Information Theoretic Security

<u>**Def**</u>: A cipher *(E,D)* over <u>(K,M,C)</u> has <u>**perfect secrecy**</u> if $\forall m_0, m_1 \in M \quad (|m_0| = |m_1|) \quad \text{and} \quad \forall c \in C$ $Pr[E(k,m_0)=c] = Pr[E(k,m_1)=c] \quad \text{where } k \stackrel{\mathbb{R}}{\leftarrow} K$

Lemma: OTP has perfect secrecy.

Proof:

$$\forall m, c: Pr \left[E(K,m)=c \right] = \frac{\# \text{Keys } K \in \mathcal{J}_{K} \text{ s.t. } E(K,m)=c}{|\mathcal{J}_{K}|}$$

Let $m \in \mathcal{M}$ and $c \in \mathcal{C}$.

How many OTP keys map $\, m \,$ to $\, c \,$?

- <u>~</u> 2
- 1
- Depends on the message

<u>Lemma</u>: OTP has perfect secrecy.

Proof:

For otp:
$$\forall m, c:$$
 if $E(k, m) = c$
 $\Rightarrow k \oplus m = c$

The bad news ...

Thm: perfect secrecy \Rightarrow $|\mathcal{H}| \geq |\mathcal{M}|$

i.e. perfect secrecy => Key-len = msg-len

- hard to use in practice !!

Review

Cipher over (K,M,C): a pair of "efficient" algs (E, D) s.t.

 $\forall m \in M, k \in K$: D(k, E(k, m)) = m

Weak ciphers: subs. cipher, Vigener, ...

A good cipher: **OTP** $M=C=K=\{0,1\}^n$

$$E(k, m) = k \oplus m$$
, $D(k, c) = k \oplus c$

Lemma: OTP has perfect secrecy (i.e. no CT only attacks)

Bad news: perfect-secrecy ⇒ key-len ≥ msg-len

Stream Ciphers: making OTP practical

idea: replace "random" key by "pseudorandom" key

1Rb is a function
$$6: \{0,1\} = \{0,1\}$$
 $n > 1$

Seed

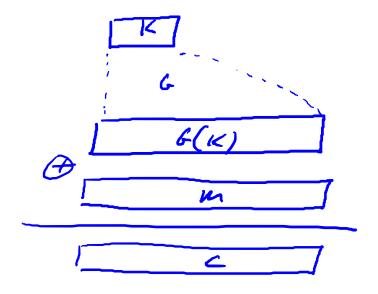
Space 1.

Left. computable by a deterministic algorithm of the sender sender.

Stream Ciphers: making OTP practical

$$C := E(K,m) = M \oplus G(K)$$

$$O(K,C) = C \oplus G(K)$$



Can a stream cipher have perfect secrecy?

- Yes, if the PRG is really "secure"
- No, there are no ciphers with perfect secrecy
- Yes, every cipher has perfect secrecy
- No, since the key is shorter than the message

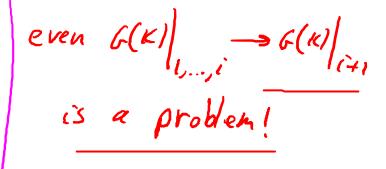
Stream Ciphers: making OTP practical

Stream ciphers cannot have perfect secrecy!!

Need a different definition of security

Security will depend on specific PRG

PRG must be unpredictable



PRG must be unpredictable

We say that $G: K \longrightarrow \{0,1\}^n$ is **predictable** if:

$$\frac{\exists \text{ "eff" alg. A and } \exists 0 \leq i \leq h-1 \leq s.t.}{\text{ for non-negligible } \{E\} \quad \{e.g. \ E=1/2^{so}\}$$

<u>Def</u>: PRG is **unpredictable** if it is not predictable

 \Rightarrow \forall i: no (eff" adv. can predict bit (i+1) for "non-neg" ϵ

Suppose G:K \rightarrow {0,1}ⁿ is such that for all k: XOR(G(k)) = 1

Is G predictable ??

Yes, given the first bit I can predict the second

No, G is unpredictable

Yes, given the first (n-1) bits I can predict the n'th bit

It depends

Weak PRGs

(do not use for crypto)

```
glibc random():

r[i] \leftarrow (r[i-3] + r[i-31]) \% 2^{32}

output r[i] >> 1
```

hever use <u>random()</u>
For crypto!!

(e.g. Kerberos V4)

Negligible and non-negligible

- In practice: ε is a scalar and
 - ε non-neg: ε ≥ 1/2³⁰ (likely to happen over 1GB of data)
 - ε negligible: ε ≤ $1/2^{80}$ (won't happen over life of key)

- In theory: ε is a function $\varepsilon: \mathbb{Z}^{\geq 0} \longrightarrow \mathbb{R}^{\geq 0}$ and
 - ε non-neg: $\exists d: ε(λ) ≥ 1/λ^d$ inf. often (ε ≥ 1/poly, for many λ)
 - ε negligible: $\forall d, \lambda \ge \lambda_d$: ε(λ) ≤ 1/λ^d (ε ≤ 1/poly, for large λ)

Few Examples

$$\varepsilon(\lambda) = 1/2^{\lambda}$$
 : negligible

$$ε(λ) = \begin{bmatrix} 1/2^λ & \text{for odd } λ \\ 1/λ^{1000} & \text{for even } λ \end{bmatrix}$$

 $\varepsilon(\lambda) = 1/\lambda^{1000}$

non-negligible

Negligible

Non-negligible

PRGs: the rigorous theory view

PRGs are "parameterized" by a security parameter λ

• **PRG** becomes "more secure" as **λ** increases

Seed lengths and output lengths grow with \(\lambda\)

For every $\lambda=1,2,3,...$ there is a different PRG G_{λ} :

$$G_{\lambda}: K_{\lambda} \rightarrow \{0,1\}^{n(\lambda)}$$

(in the lectures we will always ignore λ)

An example asymptotic definition

We say that $G_{\lambda}: K_{\lambda} \to \{0,1\}^{n(\lambda)}$ is <u>predictable</u> at position i if:

there exists a polynomial time (in λ) algorithm A s.t.

$$\Pr_{k \leftarrow K_{\lambda}} \left[A(\lambda, G_{\lambda}(k) \Big|_{1, \dots, i}) = G_{\lambda}(k) \Big|_{i+1} \right] > 1/2 + \underline{\varepsilon(\lambda)}$$

for some <u>non-negligible</u> function ε(λ)



Stream ciphers

Attacks on OTP and stream ciphers

Review

OTP: $E(k,m) = m \oplus k$, $D(k,c) = c \oplus k$

Making OTP practical using a PRG: G: $K \rightarrow \{0,1\}^n$

Stream cipher: $E(k,m) = m \oplus G(k)$, $D(k,c) = c \oplus G(k)$

Security: PRG must be unpredictable (better def in two segments)

Attack 1: two time pad is insecure!!

Never use stream cipher key more than once!!

$$C_1 \leftarrow m_1 \oplus PRG(k)$$

$$C_2 \leftarrow m_2 \oplus PRG(k)$$

Eavesdropper does:

$$C_1 \oplus C_2 \rightarrow$$

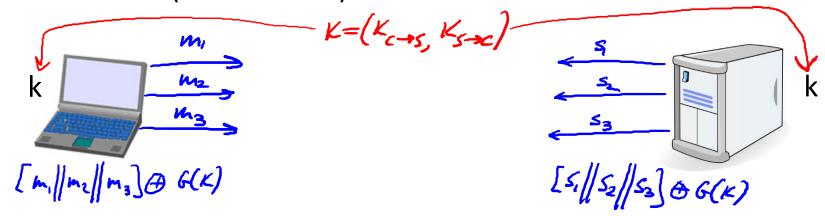
Enough redundancy in English and ASCII encoding that:

$$m_1 \oplus m_2 \rightarrow m_1, m_2$$

Real world examples

Project Venona

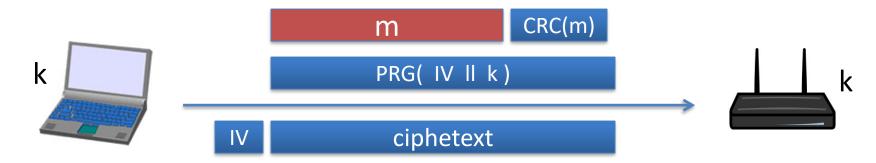
• MS-PPTP (windows NT):



Need different keys for $C \rightarrow S$ and $S \rightarrow C$

Real world examples

802.11b WEP:

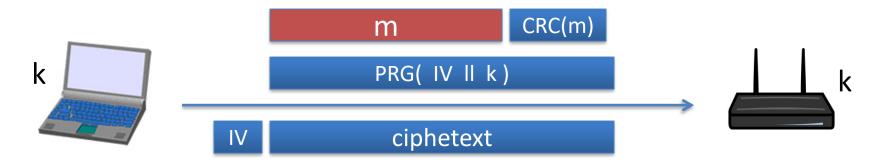


Length of IV: 24 bits

- Repeated IV after 2²⁴ ≈ 16M frames
- On some 802.11 cards: IV resets to 0 after power cycle

Avoid related keys

802.11b WEP:



key for frame #1: (1 | k)

key for frame #2: (2 | | k)

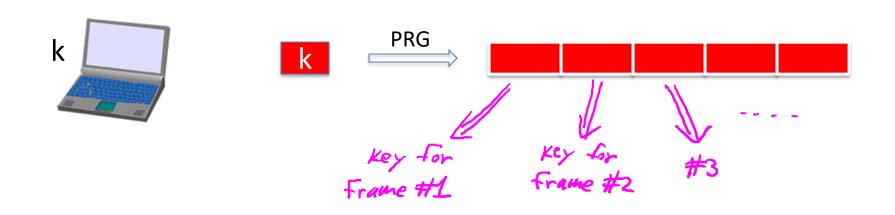
61'ts 1024 61'ts bits For the RC4 PRG:

FMS2001 => can recover 16

after 10 frames

Recent attacks = 40,000 frames

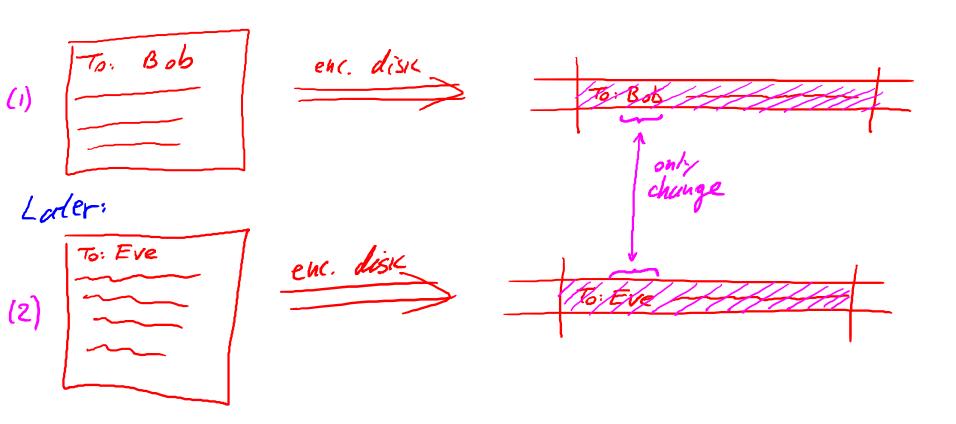
A better construction



⇒ now each frame has a pseudorandom key

better solution: use stronger encryption method (as in WPA2)

Yet another example: disk encryption



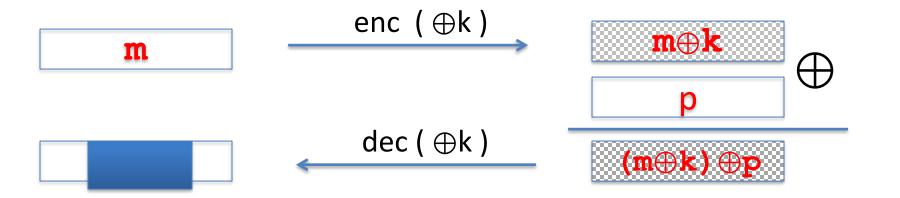
Two time pad: summary

Never use stream cipher key more than once!!

• Network traffic: negotiate new key for every session (e.g. TLS)

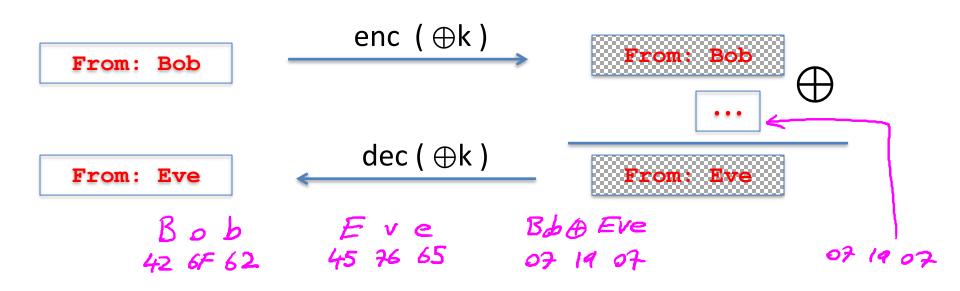
• Disk encryption: typically do not use a stream cipher

Attack 2: no integrity (OTP is malleable)



Modifications to ciphertext are undetected and have **predictable** impact on plaintext

Attack 2: no integrity (OTP is malleable)



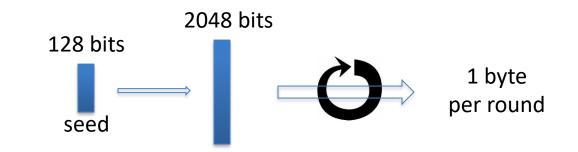
Modifications to ciphertext are undetected and have predictable impact on plaintext



Stream ciphers

Real-world Stream Ciphers

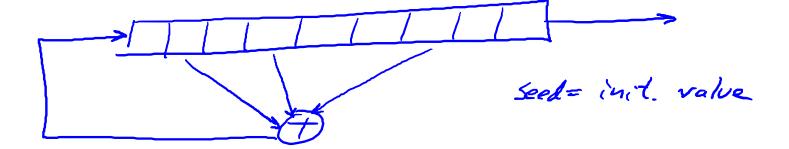
Old example (software): RC4 (1987)



- Used in HTTPS and WEP
- Weaknesses:
 - 1. Bias in initial output: $Pr[2^{nd} \text{ byte} = 0] = 2/256$
 - 2. Prob. of (0,0) is $1/256^2 + 1/256^3$
 - 3. Related key attacks

Old example (hardware): CSS (badly broken)

Linear feedback shift register (LFSR):



DVD encryption (CSS): 2 LFSRs

GSM encryption (A5/1,2): 3 LFSRs

Bluetooth (E0): 4 LFSRs

all broken

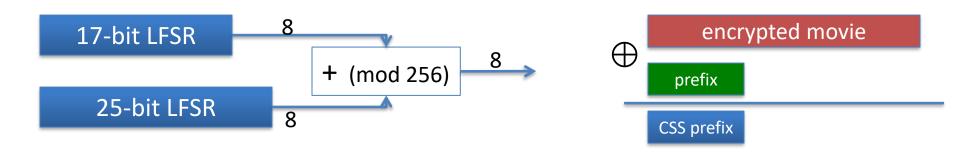
参考 教材2.2

Old example (hardware): CSS (badly broken)

CSS: seed = 5 bytes = 40 bits

Easy to break in time 22

Cryptanalysis of CSS (217 time attack)



For all possible initial settings of 17-bit LFSR do:

- Run 17-bit LFSR to get 20 bytes of output
- Subtract from CSS prefix \Rightarrow candidate 20 bytes output of 25-bit LFSR
- If consistent with 25-bit LFSR, found correct initial settings of both!!

Using key, generate entire CSS output

Modern stream ciphers: eStream

PRG:
$$\{0,1\}^s \times R \longrightarrow \{0,1\}^n$$

Seed

honce

Nonce: a non-repeating value for a given key.

$$E(k, m; r) = m \oplus PRG(k; r)$$

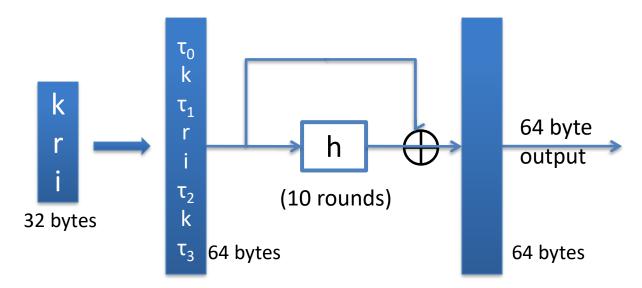
The pair (k,r) is never used more than once.

eStream: Salsa 20

Salsa20: $\{0,1\}^{128 \text{ or } 256} \times \{0,1\}^{64} \longrightarrow \{0,1\}^n$ (max n = 2⁷³ bits)

(SW+HW)

Salsa20(k;r) := H(k,(r,0)) || H(k,(r,1)) || ...



h: invertible function. designed to be fast on x86 (SSE2)

Is Salsa20 secure (unpredictable)?

Unknown: no known provably secure PRGs

In reality: no known attacks better than exhaustive search

Chacha 20

Coutinho, Murilo, and Tertuliano C. Souza Neto. "Improved linear approximations to ARX ciphers and attacks against chacha." EUROCRYPT 2021.

Performance:

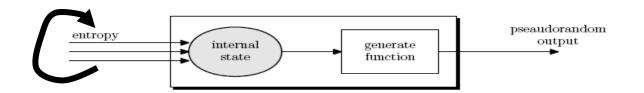
Crypto++ 5.6.0 [Wei Dai]

AMD Opteron, 2.2 GHz (Linux)

	<u>PRG</u>	Speed (MB/sec)
	RC4	126
eStream -	Salsa20/12	643
	Sosemanuk	727
	-	

Generating Randomness

(e.g. keys, IV)



Pseudo random generators in practice: (e.g. /dev/random)

- Continuously add entropy to internal state
- Entropy sources:
 - Hardware RNG: Intel RdRand inst. (Ivy Bridge). 3Gb/sec.
 - Timing: hardware interrupts (keyboard, mouse)

NIST SP 800-90: NIST approved generators



Stream ciphers

PRG Security Defs

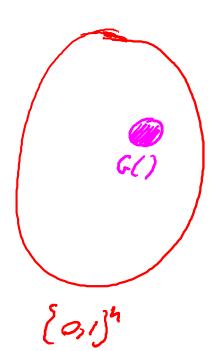
Let $G: K \longrightarrow \{0,1\}^n$ be a PRG

Goal: define what it means that



is "indistinguishable" from





Statistical Tests

not random

raudom

Statistical test on
$$\{0,1\}^n$$
:

an alg. A s.t. A(x) outputs "0" or "1"

Examples:

(1)
$$A(x)=1$$
 iff $|\#o(x)-\#1(x)| \le 10.5\pi$
(2) $A(x)=1$ iff $|\#oo(x)-\#1| \le 10.5\pi$

Statistical Tests

More examples:

(3)
$$A(x)=1$$
 iff $\max_{x} \min_{x} -o(x) < 10 \cdot \log_2(h)$

Advantage

Let G:K $\rightarrow \{0,1\}^n$ be a PRG and A a stat. test on $\{0,1\}^n$

A silly example: $A(x) = 0 \Rightarrow Adv_{PRG}[A,G] =$

Suppose G:K $\rightarrow \{0,1\}^n$ satisfies msb(G(k)) = 1 for 2/3 of keys in K

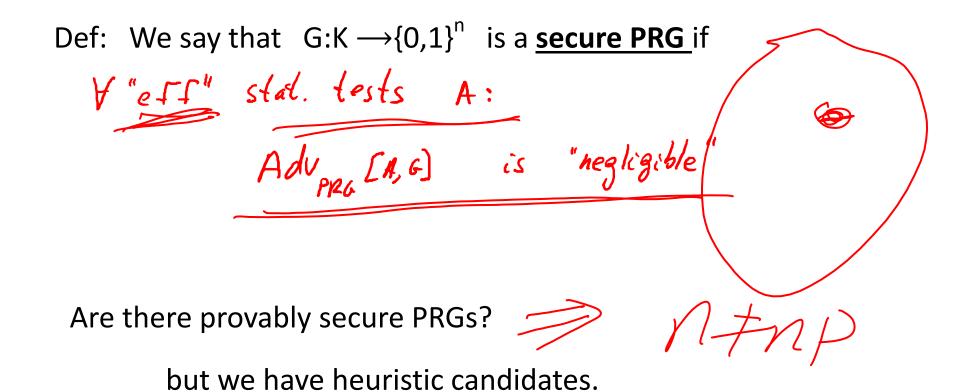
Define stat. test A(x) as:

if [msb(x)=1] output "1" else output "0"

Then

$$Adv_{PRG}[A,G] = |Pr[A(G(k))=1] - Pr[A(r)=1]| =$$

Secure PRGs: crypto definition



Easy fact: a secure PRG is unpredictable

We show: PRG predictable ⇒ PRG is insecure

Suppose A is an efficient algorithm s.t.

Easy fact: a secure PRG is unpredictable

Define statistical test B as:

$$B(x) = \begin{cases} if & A(x|_{i,...,i}) = x_{i+1} & \text{output 1} \\ else & \text{out put 0} \end{cases}$$

$$F(B(r)=i) = \frac{1}{2}$$

$$F(B(G(x))=i) = \frac{1}{2}$$

$$F(B(G(x))=i) = \frac{1}{2} + E$$

$$F(B(G(x))=i) - P(B(G(x))=i) \gg E$$

Thm (Yao'82): an unpredictable PRG is secure

Let $G:K \longrightarrow \{0,1\}^n$ be PRG

"Thm": if \forall i \in {0, ..., n-1} PRG G is unpredictable at pos. i then G is a secure PRG.

If next-bit predictors cannot distinguish G from random then no statistical test can!!

Let $G:K \longrightarrow \{0,1\}^n$ be a PRG such that from the last n/2 bits of G(k)it is easy to compute the first n/2 bits.

Is G predictable for some $i \in \{0, ..., n-1\}$?

No



Stream ciphers

Semantic security

Goal: secure PRG ⇒ "secure" stream cipher

Review

Stream Cipher: making OTP practical using a PRG

Linear feedback shift register(LFSR)

Bad constructions: CSS, 802.11b WEP, MS-PPTP

Good construction: Salsa20

Security def. of PRG: unpredictable <=> security [no "effi" sta. test]

Notation

Let P_1 and P_2 be two distributions over $\{0,1\}^n$

Def: We say that P_1 and P_2 are

computationally indistinguishable (denoted $\Re \approx \Re$)

if
$$\forall$$
 "eff" stat. tests A

$$|\Pr[A(x)=1] - \Pr[A(x)=1]| < \text{negligible}$$

$$|x \leftarrow P_1| = ||x - P_2|| = ||$$

Example: a PRG is secure if $\{k \leftarrow^R K : G(k)\} \approx_p uniform(\{0,1\}^n)$

What is a secure cipher?

Attacker's abilities: **obtains one ciphertext** (for now)

Possible security requirements:

attempt #1: attacker cannot recover secret key

$$E(K,M)=M$$

attempt #2: attacker cannot recover all of plaintext $E(\kappa, m_o || m_i) = m_o || m_i \oplus \kappa$

Recall Shannon's idea:

CT should reveal no "info" about PT

Recall Shannon's perfect secrecy

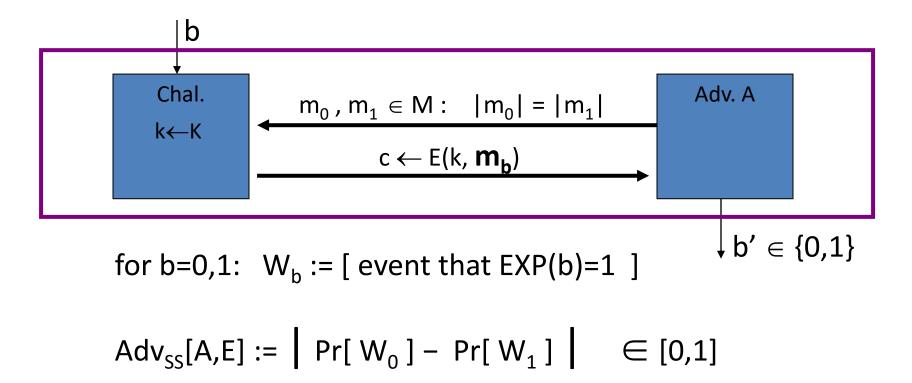
Let (E,D) be a cipher over (K,M,C)

```
(E,D) has perfect secrecy if \forall m_0, m_1 \in M (|m_0| = |m_1|)  \{E(k,m_0)\} = \{E(k,m_1)\} \text{ where } k \leftarrow K  (E,D) has perfect secrecy if \forall m_0, m_1 \in M (|m_0| = |m_1|)  \{E(k,m_0)\} \approx_p \{E(k,m_1)\} \text{ where } k \leftarrow K
```

... but also need adversary to exhibit $m_0, m_1 \in M$ explicitly

Semantic Security (one-time key)

For b=0,1 define experiments EXP(0) and EXP(1) as:



Semantic Security (one-time key)

Def: \mathbb{E} is **semantically secure** if for all efficient A

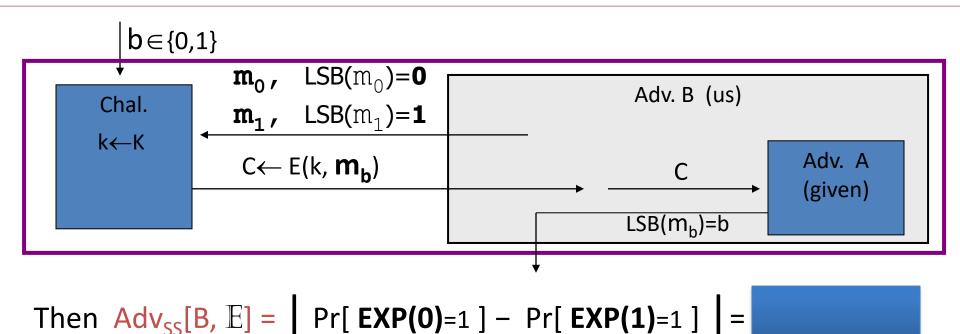
 $Adv_{SS}[A,E]$ is negligible.

 \Rightarrow for all explicit m_0 , $m_1 \in M$: $\{E(k,m_0)\} \approx_p \{E(k,m_1)\}$

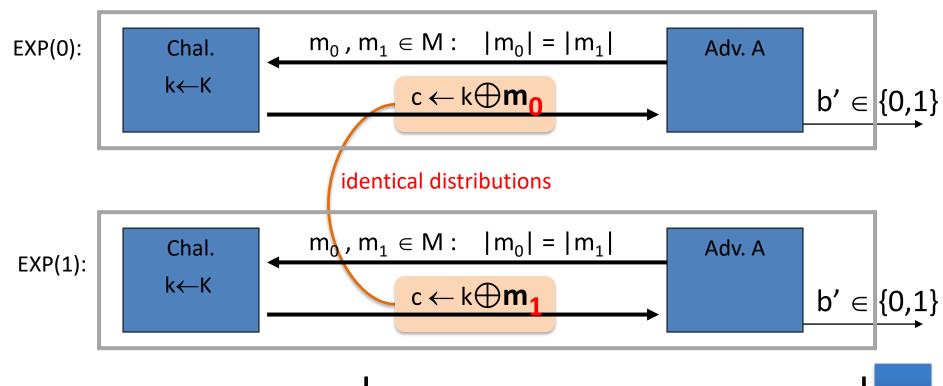
Examples

Suppose efficient A can always deduce LSB of PT from CT.

 \Rightarrow \mathbb{E} = (E,D) is not semantically secure.



OTP is semantically secure



For <u>all</u> A: $Adv_{SS}[A,OTP] = | Pr[A(k \oplus m_0)=1] - Pr[A(k \oplus m_1)=1] |$



Stream ciphers

Stream ciphers are semantically secure

Goal: secure PRG ⇒ semantically secure stream cipher

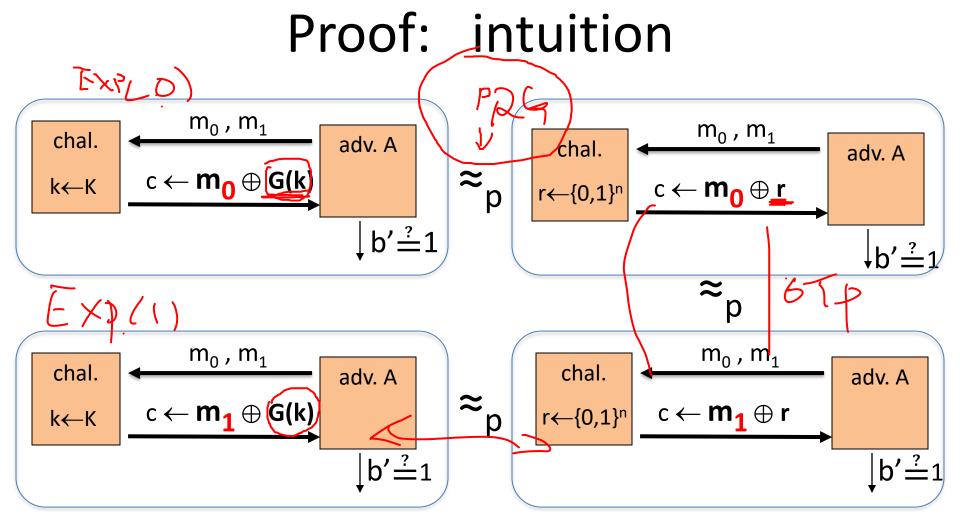
Stream ciphers are semantically secure

Thm: $G:K \longrightarrow \{0,1\}^n$ is a secure PRG \Rightarrow

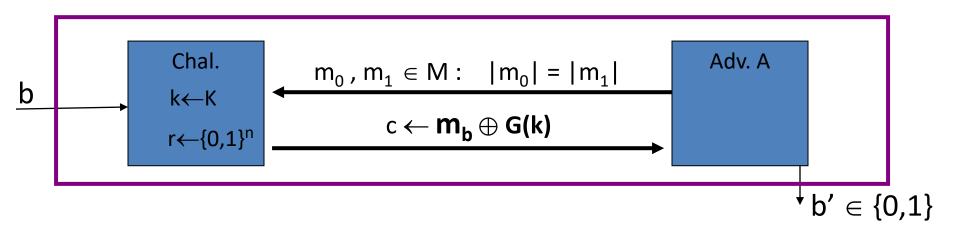
stream cipher E derived from G is sem. sec.

∀ sem. sec. adversary A , ∃a PRG adversary B s.t.

$$Adv_{SS}[A,E] \leq 2 \cdot Adv_{PRG}[B,G]$$



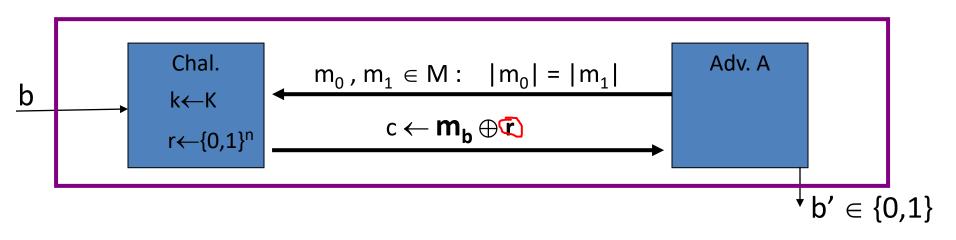
Proof: Let A be a sem. sec. adversary.



For b=0,1: $W_b := [event that b'=1].$

$$Adv_{SS}[A,E] = | Pr[W_0] - Pr[W_1] |$$

Proof: Let A be a sem. sec. adversary.



For
$$b=0,1$$
: $W_b := [event that b'=1].$

$$Adv_{SS}[A,E] = | Pr[W_0] - Pr[W_1] |$$

For b=0,1:
$$R_b$$
:= [event that b'=1]

Proof: Let A be a sem. sec. adversary.

Claim 1:
$$\frac{\left|\Pr[R_0] - \Pr[R_1]\right| = Adv_{ss}[A, otp] = 0}{\exists B: \left|\Pr[W_b] - \Pr[R_b]\right| = Adv_{pre}[B, e]} \qquad \text{for } b = g \text{ } l$$

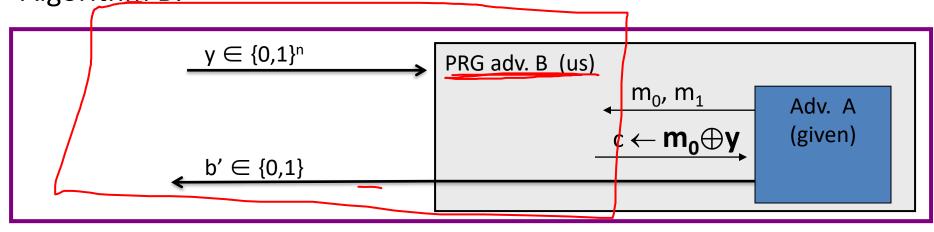
$$0 \qquad \qquad \Pr[W_0] \qquad \Pr[R_b] \qquad \Pr[W_1] \qquad \qquad 1$$

$$Adv_{pre}[B, e] \qquad Adv_{pre}[B, e] \qquad Adv_{pre}[B, e]$$

$$\Rightarrow$$
 Adv_{SS}[A,E] = $|Pr[W_0] - Pr[W_1]| \le 2 \cdot Adv_{PRG}[B,G]$

Proof of claim 2:
$$\exists B: Pr[W_0] - Pr[R_0] = Adv_{PRG}[B,G]$$

Algorithm B:



$$\underline{Adv_{PRG}[B,G]} = \begin{cases} Pr & \left[B(r) = 1 \right] - Pr \left[B(h(k)) = 1 \right] \\ r \in \left[a_{1} \right]^{n} \left[\frac{B(r) = 1}{k} \right] - Pr \left[\frac{B(h(k))}{k} \right] \end{cases} = Pr \left[\frac{Pr}{N_{0}} \right] - Pr \left[\frac{N_{0}}{N_{0}} \right]$$

End of Segment