

# Project Venona — wikipedia

14 Jan 2025

## Information theoretic security

### Shannon Security

$$\forall m, c: \Pr(\mathcal{P} = m \mid \mathcal{C} = c) = \Pr(\mathcal{P} = m)$$

Equivalent

### Perfect Indistinguishability

- $m_1, m_2$  picked by Adv.
- one of them randomly encrypted by the challenger
- Adv's advantage in distinguishing which of the two messages produced  $c = 0$

$$\forall m_1, m_2, \forall c$$

$$\Pr(\mathcal{P} = m_1 \mid \mathcal{C} = c) = \Pr(\mathcal{P} = m_2 \mid \mathcal{C} = c)$$

① Shannon Security  $\Rightarrow$  Perfect indistinguishability

$$\Pr(\mathcal{P} = m_1 \mid \mathcal{C} = c) = \Pr(\mathcal{P} = m_1) \text{ for any } m_1, c$$

(defn of conditional prob)

$$\frac{\Pr(\mathcal{P} = m_1 \cap \text{Enc}(k, m_1) = c)}{\Pr(\mathcal{C} = c)} = \Pr(\mathcal{P} = m_1)$$

$$\frac{\cancel{\Pr(\mathcal{P} = m_1)} \cdot \Pr(\text{Enc}(k, m_1) = c)}{\Pr(\mathcal{C} = c)} = \cancel{\Pr(\mathcal{P} = m_1)}$$

$$\Pr_{k, c}(\text{Enc}(k, m_1) = c) = \Pr_c(\mathcal{C} = c)$$

$$= \Pr(\text{Enc}(k, m_2) = c)$$

$\Rightarrow$

$$P_c(\text{Enc}(k, m_1) = c) = P_c(\text{Enc}(k, m_2) = c)$$

(II)

Perfect indist  $\Rightarrow$  Shannon security

This is the strongest notion of encryption security  
— even against a exponential time attacker,  
this security is enough

Practical instantiation . OTP

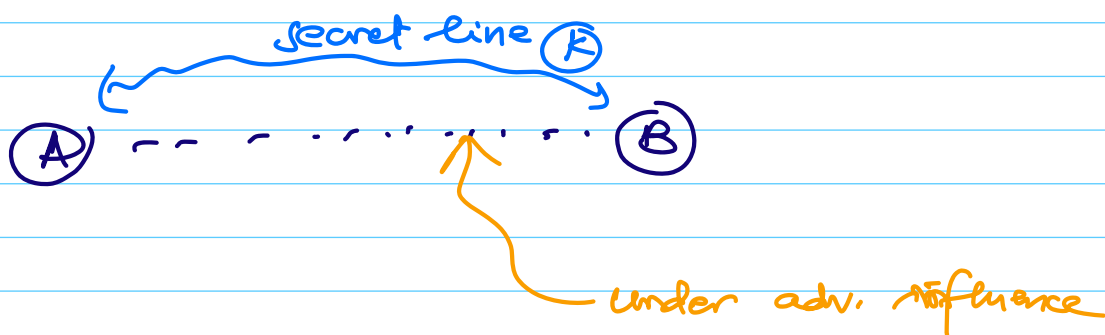
$$\text{Enc}(k, m) = k \oplus m$$

$$\text{Dec}(k, c) = k \oplus c$$

$$k, m, c \in \{0, 1\}^n, \quad k \text{ was chosen uniformly at random}$$

Practical problem -

to encrypt a message of length  $n$  bits,  
we need a secret key of length  $n$  bits



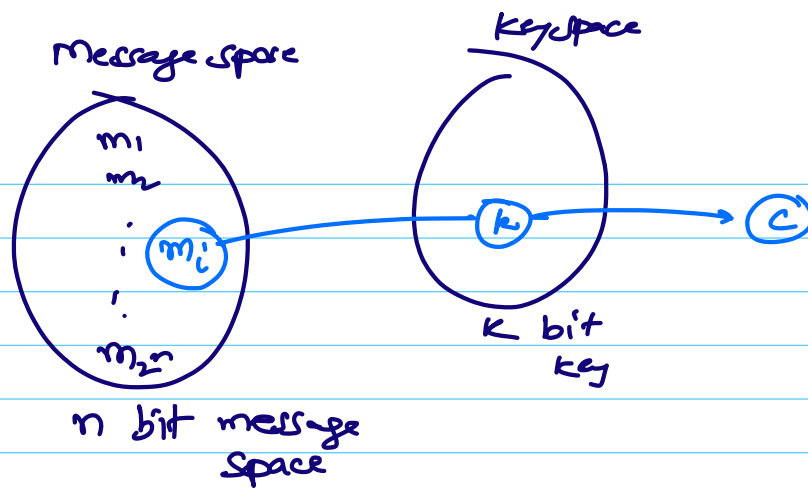
Que: Can we reduce randomness requirement?

i.e message length =  $n$  bits

but key length =  $k$  bits where  $k < n$

Claim:

In this setting, perfect security is impossible



Attacker's idea:

- Pick the ciphertext  $c$
- Decrypt by all possible keys

$$\mathcal{P}' = \{ \text{Dec}(k_1, c), \text{Dec}(k_2, c), \text{Dec}(k_3, c), \dots \}$$

$$|\mathcal{P}'| \leq |K|$$

But

$$|K| < |P|$$

$\Rightarrow \exists m^* \in \mathcal{P}$  which is not in  $\mathcal{P}'$

$$\Pr(\mathcal{D} = m^* | C = c) = 0$$

But for some message  $m' \in \mathcal{P}'$

$$\Pr(\mathcal{D} = m' | C = c) \neq 0$$

$\Rightarrow$  Practical limitation in using OTP in real-life

## Computational Security

(this is weaker than perfect security)

Ⓘ Now we will limit the power of the adversary.

- Adv can not run for exponential time
- allowed only poly. time operations
- Security parameter =  $n$

Ⓜ Attacker will have some non-zero prob. of winning the distinguishability game

$$P_0(\dots) - P_1(\dots) \neq 0$$

Now

$$P_0(\dots) - P_1(\dots) = \epsilon$$

$\downarrow$   
negligible in  $n$   
 $\approx \frac{1}{2^n}$

Negligible function:

Pr of winning the game should be "negligible"

$$\text{poly}(n) \times (\text{Prob of win}) = \text{negl}(n)$$

$$\Rightarrow \text{negl}(n) = \frac{\text{poly}(n)}{\text{exp}(n)} \quad \text{or something that sort}$$

$$n \rightarrow \infty, \quad \text{negl}(n) \rightarrow 0$$

# How to create a computationally secure encryption scheme?

Why did OTP kill any statistics in the input?

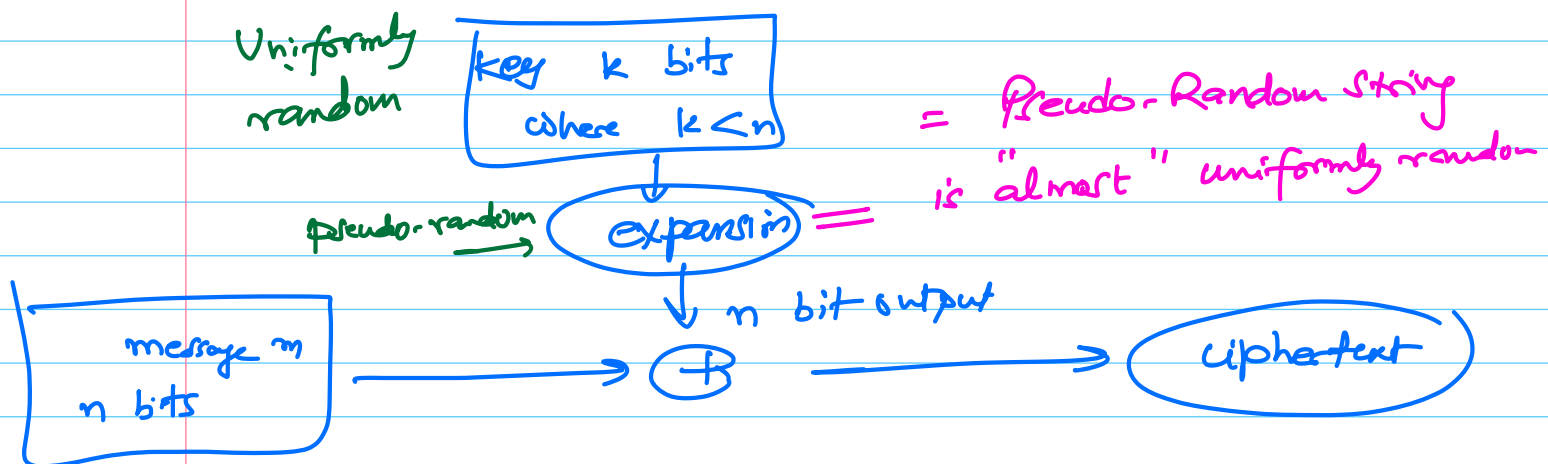
$$\Pr(b=0) = 0.9 \quad , \quad \Pr(b=1) = 0.1$$

$$\Pr(k=0) = \Pr(k=1) = 0.5 \quad \text{uniformly random key bit}$$

Que

$$\Pr(b \oplus k = 0) = ?$$

$$\begin{aligned} &= \Pr(\underline{b=k=0}) + \Pr(\underline{b=k=1}) \\ &= \frac{1}{2} \times 0.9 + \frac{1}{2} \times 0.1 \\ &= \frac{1}{2} \end{aligned}$$




Example: `rand()` function in C, Java, ...

What is a "random" string?

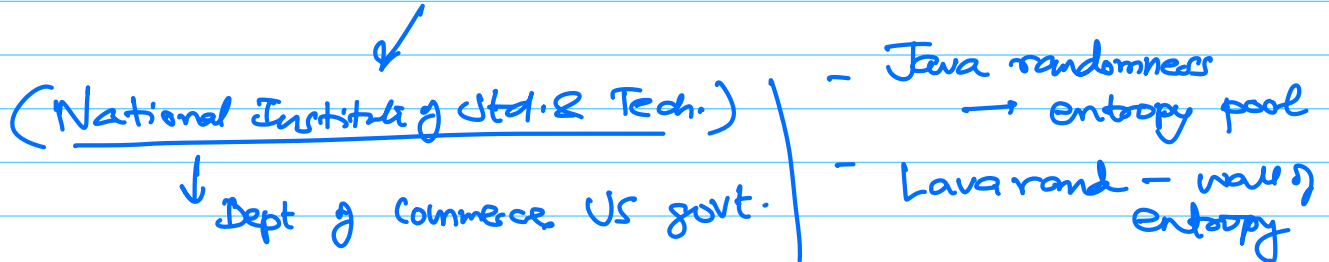
1. is 000...0 random?
2. is 110011... random?
3. is 011001... random?

Source is random  
(Not the strings)

Yao - given a string  (Next bit test)

## Statistical test of randomness

NIST Test suite



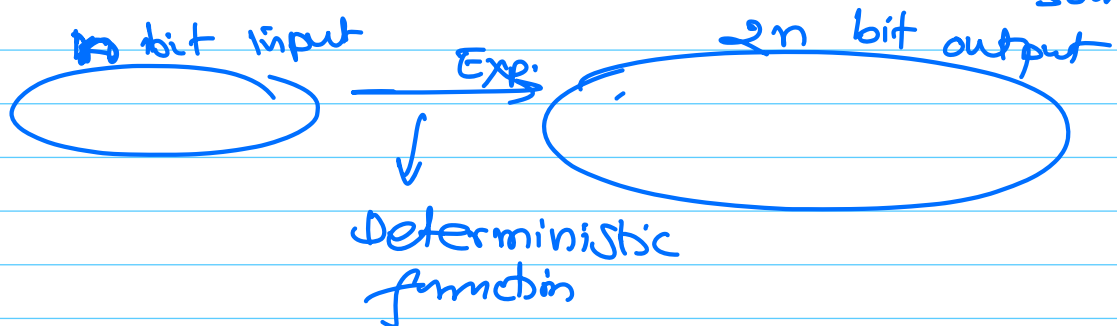
## PRG (Pseudo-Random Generator)

Idea:

input: small amount of randomness

output: large " of pseudo-randomness

Something which could be identified by "someone" as only from random vs. pseudo-random source



How many possible outputs for this function?

$$= 2^n$$

$$\text{Size of output set} = 2^{2n}$$

if we pick any random string in the output set, it is not likely to have been produced by this algo

$$\Pr. \left( \begin{array}{l} \text{a random string} \\ \text{from output is} \\ \text{actually produced} \\ \text{from this algo} \end{array} \right) = \frac{1}{2^n}$$

Brute force attack will win with prob  $\approx 1$   
— to prevent this, we will not  
allow attacker this exponential power

16-Jan-25

PRG - Pseudo-random generator ( $G$ )

Idea: a PRG takes a short seed and produces a long output which "looks like" random  
2. PRG is a deterministic function.

$$G: \{0,1\}^s \rightarrow \{0,1\}^l \quad \text{where } \frac{\text{expansion}}{l > s}$$

↓  
should be "looking like" random

Adv A : PPT : Probabilistic Polynomial Time algorithm

↙ (i) internal coin tosses  
(ii) also gets a random string as input

Real world  
 $s \leftarrow \{0,1\}^n$   
 $y = G(s)$

Ideal world  
 $y \leftarrow \{0,1\}^l$