

Intro to Modern Cryptography

- J. KATZ + Y. LINDELL

POTS

FANTASTIC Q When is a problem Infosec one? [Ans] When the problem is impossible to solve logically / perfectly.

Will be shown with standard examples.

e.g.: Hashing password.

Theoretically impossible to be perfect if length is not infinite.

e.g.: Secure communication

$$\text{① } \text{to } \text{info}(R) = \text{info}(eve) \quad (S) \xrightarrow{\text{Enc}} (R)$$

$$\text{② } \text{to } \text{info-rec}(R) = \text{info-rec}(eve)$$

e.g.: Data integrity

If m was sent modified to m' \rightarrow It is the same to receiver.
and if $m' \neq m$, L cannot thus identify.



OO Eg of non infosec \rightarrow Infosec

• problem in distributed computing \therefore now an infosec problem.

Solution: use signatures \Rightarrow implying signatures are impossible

FASCINATING Q. How to logically solve/circumvent a logical impossibility?

[Ans] Bring in another impossibility and make it destructively interfere with the original one.

We focus on 4-5 sources of impossibilities in the semester.

Course: See impossibilities

Introduce others

Solve them

1 per month approx.

Sources of Impossibility.

① Computational Hardness [Resource Complexity]

Only happy
cats are secure.

② Practical Uncertainties

Speed of light
distance stuff

③ Natural Limits

④ Logical / Philosophical Impossibilities

Random Words

- Hamming Distance

- Information security is God
- all non-trivial works of
science must include
Info sec

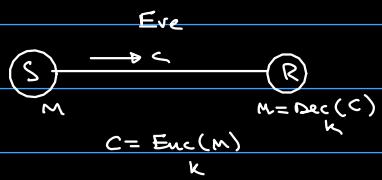
- logical norgo

Topics to cover

- Kerckhoff's Principle
- Designing/Breaking classical ciphers.

Starting off with secure communication networks.

- traditional ciphers, and how to break them.
- Shannon next class.
 - defined information
 - path breaking.

Caesar Cipher

$$c = (m + 3) \% n_c$$

m = message

c = coded message

n_c = no. of characters in alphabet.

Big talk about his perspective of infosec as an 'art', and a bigger rant on what is art and what is science.

Exact words in book.

Kerckhoff's Principle

Security of a system must NOT depend on the OBSCURITY of the algorithm, rather must solely depend on the SECRECY of the KEY.

Kerckhoff's Reasons

1. Algorithms are reverse engineerable.

$$\text{eq: } h(x) = bx + c$$

Attacker can feed x_1, x_2, \dots, x_n and see that all outputs $h(x_1), \dots, h(x_n)$ for $g | h(x_i) - h(x_j)$. And then solve for c .

2. Updation/ Recovery Complexity.

if passwd rarely in secure system: change pass.
if algo "in obscurity" // fixed.

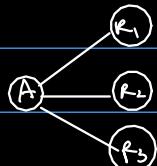
3. Secure Memory is costly.

ASK ACHIEV'A

Prob in modern times: UTM is algo and full input is key.

— bad information storage efficiency.

4. Scalable



Without: Diff algorithm for everyone.

With: only 'key' changes among people.

Additional Reasoning

1. Ethical Hacking

hypothesis: no system is secure.

bug (algo) will be found because \exists L

→ bug exists
→ nonethical people exist
→ only " " search for bug.
→ " " will find "
→ bug F, take that L

2. Standards

needed for efficiency.
if an algo is used by every system, it should obviously follow

Thus we can see why Caesar cipher fails.

Next iteration:

Shift Cipher:

$$c = (x+k) \bmod n_k$$

m = message

c = coded message

k = key

n_c = no. of characters in alphabet.

ATTACK

1. Can be broken by humans

brute force
→ If keyspace is smal, attackez.

first principle learnt!

Principle of large key space

how do we know

the key is

correct

2. Autobreaking:

- frequency analysis

$$p_i = P(i^{\text{th}} \text{ char in } m)$$

$$\text{Precompute } \sum_{i=0}^{25} p_i^2 \approx 0.065$$

$$q_i = P(" " = c)$$

Now compute

$$\sum_{i=0}^{25} (p_i q_{i+k}) \begin{array}{l} \xrightarrow{\text{wrong guess}} \\ \xrightarrow{\text{correct guess}} \end{array} \frac{1}{26} \approx 0.065$$

Next iteration:

Monoalphabetic Substitution Cipher

- Diff alphabets shift by different amounts.

- no repetitions allowed

- basically, permutation

applying first principle

for brute force: $26!$ keys to search

Attack

$$\forall i \exists j : q_j \approx p_i$$

⇒ 1. Sort q_i 's

2. Sort p_i 's

since distribution is same,

$$\text{e.g.: } P_a = Q_n, P_c = Q_t, P_t = Q_b$$

Issue: susceptible to frequency attacks.

Next iteration

Vigenère Cipher

i. does not maintain frequency

e.g.: $\begin{array}{l} \text{h e l l o} \\ \text{s e a s e} \\ \text{z i l d} \end{array}$

FAILED ATTACKS

broke for: too many keys

• freq anal: freq. not maintained

Can be broken if:

1. key length known

2. k " is findable.

ATTACK

PART 1:

if we know length of key,

e.g.: 3

$c_0 c_3 c_6 \dots$

partition ciphertext into (length) parts

$c_1 c_4 c_7 \dots$

then shift cipher attack on all (length) parts.

$c_2 c_5 c_8 \dots$

PART 2:

Guess an L .

take a string $c_0 c_1 c_2 \dots$

$$\text{check if } \sum_{i=0}^{25} q_i^L = \sum_{i=0}^{25} p_i^L$$

easy

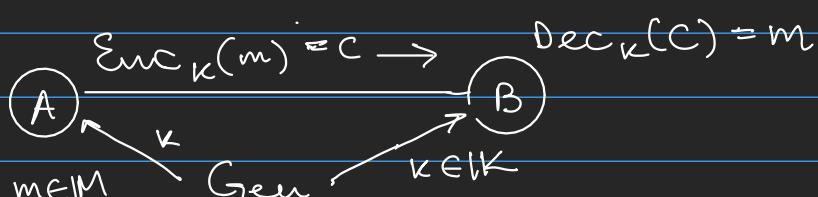
Topics for 10/01/2020

- Shannon's Perfect Secrecy
- Vernam Cipher is perfect (one-time pad)
- Limitations of Shannon's Approach

* Shannon's Perfect Secrecy

Definition of perfectly secure cipher.

An encryption scheme is a 4-tuple
 $\langle \text{Gen}, \text{Enc}, \text{Dec}, \mathcal{M} \rangle$



$$\text{Dec}_K(\text{Enc}_K(m)) = m$$

Perfectly secret
encryption
scheme.

Schemes that meet this
are largely impractical

An encryption scheme is said to be perfectly secret if for all probability distributions over \mathcal{M} , and for all $m \in \mathcal{M}$, for all $c \in \mathcal{C}$ [where $P(C = c) > 0$],

$$P[\text{Message} = m | \text{Ciphertext} = c] = P[\text{Message} = m]$$

We will prove one time pad is perfectly secret.

$$P[M = m | C = c] = P[M = m]$$

Random Variables

Vernam Cipher

$$\mathcal{M} = \mathcal{K} = \mathcal{C} = \{0, 1\}^n \quad (\text{n-bit space})$$

$$\text{Gen : } K \leftarrow_R \{0, 1\}^n$$

$$\text{Enc : } C = m \oplus k$$

$$\text{Dec : } m = c \oplus k$$

- correctness is fairly obvious and not shown here.

$$\begin{aligned} a \oplus a &= 0 \\ m \oplus a \oplus a &= m \oplus 0 = m \\ \text{Dec}_k(\text{Enc}_k(m)) &= m \oplus k \oplus k = m \\ &\quad (\text{Proved}) \end{aligned}$$

PROOF

First Showing that definition of perfect security is equivalent to

$\forall p \in \mathcal{P}$ over \mathcal{M}

$\forall m \in \mathcal{M}, \forall c \in \mathcal{C}$

$$P[C = c | M = m] = P[C = c]$$

have to do an iff.

Suppose this holds.

$$P[C = c | M = m] = P[C = c]$$

multiply both sides,

$$\frac{P[M = m]}{P[C = c]} \cdot P[C = c | M = m] = P[M = m]$$

$$\Rightarrow P[M = m | C = c] = P[M = m]$$

Second

Showing $\forall c \in \mathcal{C}, \forall m_0, m_1 \in \mathcal{M}$ is eq to

$$P[C = c | M = m_1] = P[C = c | M = m_0]$$

2 \rightarrow 1 is trivial ($P(C = c | M = m) = P(C = c), \forall m = m_0, m_1$)

$$P[C = c] = \sum_{m \in \mathcal{M}} P[C = c | M = m] P[M = m]$$

$$P$$

$$= P \sum_{m \in \mathcal{M}} P[M = m] = P$$

Workability $\propto \frac{1}{\text{Intuitiveness}}$

The above probability tells us that the encryption of m_0 and m_1 are indistinguishable.

For Vernam Cipher:

$$\text{LHS} = P[C = c \mid M = m_0]$$

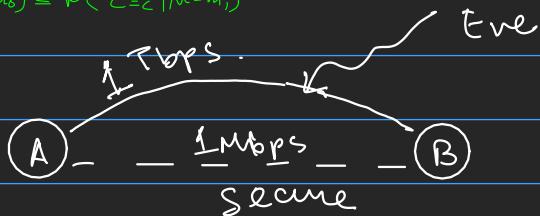
((intuition))

$$= P[C = c \oplus k] = P[k = c \oplus m_0]$$

equivalence

$$= \frac{1}{2^n} = \text{RHS} \quad (P \text{ is not dependent on } m_0)$$

Now also for m_1 ,
then by $P(C=c \mid M=m_0) = P(C=c \mid M=m_1)$



Uses of Vernam Cipher:

- i) To save your shady business from the feds, encrypt the data and delete after raid.
- ii) Use secure channel on low load days to transfer keys and on high load days, use insecure channel by encrypting.

Now, bifurcation in the field.

Symmetric key cryptography.

Two impossibilities

Slow secure channel
+ fast insecure channel

Impossible if perfect security.

Public Key cryptography

Only in secure channel \Rightarrow slow secure channel

Both are impossible

Now showing that limitations of Vernam cipher apply to any perfect cipher system as per Shannon's definition.

Universality of Shannon theorem.

Thm: For any perfectly secret encryption scheme $|K| \geq |M|$

\Rightarrow no. of bits required to store message is lower bound to no. of bits in key.

Shannon did:

$$H(|K|) \geq H(|M|)$$

We use a handy bypass to this.

ASK ATHREYA FOR INTUITION.

Missing here:

K is compressible.

Proof:

Suppose not.

(the contrary)

$$|K| < |M|$$

some ciphertext C

we will show this directly

implies this cannot be perfectly secret

$$D = \{m \mid \exists k \in K \text{ } Dec_k(c) = m\}$$

all m 's s.t. there is a key mapping it to c

$$\text{now, } |D| \leq |K| \therefore < |M|$$

$$\Rightarrow \exists m^* \in M \text{ s.t. } m^* \notin D$$

$\log |K|$ is the max entropy of $|K|$.

and $H(K) = \log |K|$ for uniform dist $|K|$

$H(M)$ can never be larger than

so we have to consider a dist where $P(M=m^*) \neq 0$

$$\therefore P[M=m^* \mid C=c] = 0$$

$$\log(|M|)$$

if $|K| < |M|$,

$$\log |K| < \log |M|$$

not possible.

\Rightarrow Scheme is not perfectly secret

\Rightarrow For perfectly secret scheme, $|K|$ must be at least $|M|$.

∴ one-time pad not a one-off.

17.1.20

Oh no it's Chiranjeevi

Class on either 1. Finite Fields

2. Elliptic Curve.

Groups:	(set, binary operation) satisfying axioms	eg: $(\mathbb{Z}, +)$
- closure	- Identity	
- associative	- Inverse	

for group G ,

if $H \subset G$ and satisfies property, H is a subgroup of G .

Cyclic group if $a \in G$, and $G = \{a^0, a^1, a^2, \dots\}$

eg. of $(\mathbb{Z}_n, +)$ groups

Next:

Ring, Integral Domain, Field.

Ring

$(R, +, \cdot)$ two binary operations on a set R

a) $(R, +)$ is a commutative group

b) Closure: $a, b \in R \Rightarrow a+b \in R$

c) Associative: $(a+b)+c = a+(b+c) \forall a, b, c \in R$

d) Distributive laws: $(a+b).c = a.c + b.c ; a.(b+c) = a.b + a.c$

if $a.b = b.a \forall a, b \in R$

Review

1940s

→ Shannon's Perfect Secrecy
and Pessimistic Result

$$[\text{P.S.} \Rightarrow |\mathcal{K}| \geq |\mathcal{M}|]$$

1970s - early 80s

- Two famous Relaxations

One Universal
Assumption

Class focuses on Mid 70s - early 80s.

They noted that due to pessimistic result, they had to relax some fundamental assumptions, and hopefully they would be sufficient.

Current situation: The two necessary relaxations are almost sufficient, assuming "some yet undiscovered object exists"

In Perfect Secrecy

Adversary was unbound

However, we know [adversary is computationally bounded] → First Relaxation

Zero error

with finite passwords, one must necessarily work.

[Instead of zero error, "small negligible value of err"] → Second Relaxation

turns out, these 2 are almost sufficient: Security against

- computationally bounded adversary
- negligible error

Definition of efficiency / practicality: bounded by polynomial time

↳ helps model adversary as Probabilistic Polynomial Turing Machines (PPTM)

What is a
"computationally bounded
adversary"?

Security Parameter k

Practical Adversaries: PPTM in Security Parameter

Negligible Function: A function $\mu(n)$ is said to be negligible $\xrightarrow{n \rightarrow \infty}$ if it is smallerthan the inverse of any polynomial for a sufficiently large n .

$$\left\{ \forall p() \quad \exists n_0 \text{ s.t. } \forall n \geq n_0, \mu(n) \leq \frac{1}{p(n)} \right\}$$

⇒ Secrecy (Relaxed): If PPTM Adversary; $P[A \text{ can break } h] \leq \text{negl}()$

Next

define
"negligible error"

What exactly do we need to reasonable negligibility - and very few functions do not work.

Adversary trying to crack system will try multiple times. } probability grows polynomially.
 ↳ will try polynomial no. of times. }
 now, $\frac{1}{\text{fixed no.}}$ will be surpassed in time.
 We need $\frac{1}{\text{polynomial limit + 1}}$ negligibility.

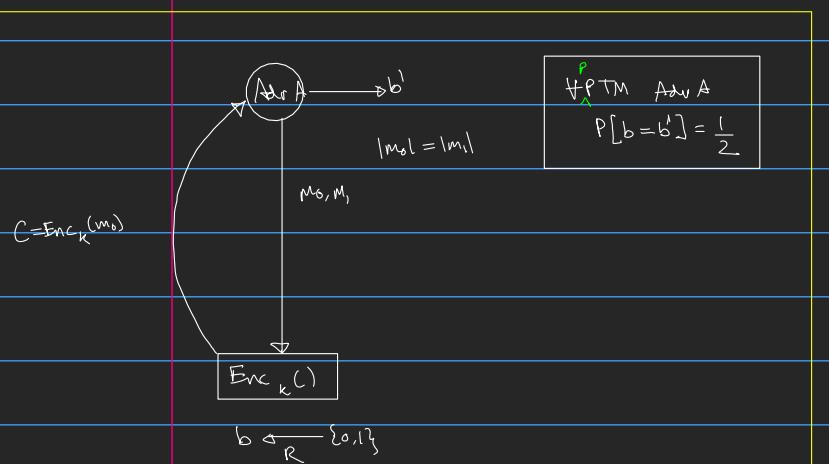
negligible means "smaller than the inverse of any polynomial".
 for example exponential growth

$\nexists p()$, $\exists n \text{ s.t. } \forall n \geq n_0$

$$\frac{1}{2^n} \leq \frac{1}{p(n)}$$

} future proofs, in the sense:

- with increased computation (polynomial bound), we keep the algorithm, increase security parameter n only.



Equivalent Definition of Perfect Secrecy.

An encryption scheme is ciphertext-only secure if
 $\forall \text{PTM } A, P[b = b'] \leq \frac{1}{2} + \text{negl}(n)$
 (P of selecting correct message from the message space is negligibly more than random)

other than perfect secrecy.

Bottleneck: no known encryption scheme is known to meet this definition.

Λ

Apart from these 2 relaxations, if _____ exists, it is possible.

- pseudorandom generator
- collision resistant hashing
- one way hash functions exist

All are the same mathematical objects. One way functions.

Does this exist? No fun due there is a faster.

Field, at this point, trifurcated.

① Heuristic Security.

(Secure because we have assumed it is)

e.g.: 1960s - mid 1990s was DES

late 1990s - current is AES

} focusing least attention to this.

because

- heuristics are prone to obsolescence
- no answer for security

③ Proven Security $\rightarrow \emptyset$ | taught when it exists

② Provable Security

We will give a conditional proof.

Definition of One-Way Function

A function $f: \{0,1\}^* \rightarrow \{0,1\}^*$ is said to be one way if : (easy to compute, hard to invert)

a) Easy to compute, $x \rightarrow f(x)$, polynomial time.

b) Hard to invert : $\nexists \text{PPTM } A$

$$P[A(f(x)) = y \mid f(x) = f(y)] \leq \text{negl}(|x|)$$

Weaker : $P[A(f(x)) = x] \leq \text{negl}(|x|)$; assumes $f^{-1}(x)$ exists

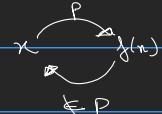
$P(\text{successfully inverting function})$ is negligible in previously defined terms.

Trio of impossibility based: Two famous relations + one way functions exist.

Not too sure
on topic heading.
Ask someone.

Now, demonstrating that proving that one way functions exist is at least NP-hard.

One Way functions:



NP, complexity theory

proving that reverse process has to be NP,

$L \in \text{NP}$ if L has an efficient verifier

decider: given x , $x \in L$?

Verifier: given proof, we can verify L .

if $x \in L$, \exists at least 1 v which works

A verifier V is said to verify L if

$$L = \{x \mid \exists v, V(x, v) = \text{accepts}\}$$

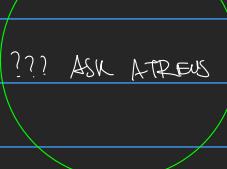
Now,

if \exists p(t) NDTM, \exists efficient verifier

$\nexists \text{PPTM, not in BPP}$
 $\text{and } P \subseteq \text{BPP} \subseteq \text{NP}$
 $\Rightarrow \text{not in P}$

Reverse process has efficient verifier

as it must be NP



TODAY

- Pseudo-randomness
- An encryption scheme
- Discrete Log Problem



24.01.2020

Discrete Logarithm Problem (DLP)

Given the group \mathbb{Z}_p^* and its generator g , and $y = g^x$ Find x .

p is exponential in terms of $\log p$
bits of input.

there are no known algorithms that can solve this in polynomial time.

conjecture: this is a one-way function

Concept of one-time pads using computational entropy.



Pseudo-Random number generator

An efficient deterministic code G , that inputs n -bit string and outputs

$\lambda(n)$ -bit string is a PRG if

a) Expansion: $\lambda(n) > n$

b) Pseudo-randomness: If PPTM distinguishes D (that try to distinguish pseudo-randomness from randomness)

$$P[D(u_{\text{new}}) = 1] - P[D(G(u_n)) = 1] \leq \text{negl}(n)$$

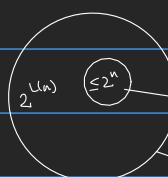
exist if one-way functions exist

G is a PRG if no efficient distinguisher can distinguish
between a uniformly distributed random of length $\lambda(n)$
and a expanded number from a smaller key of length n



Use case: send key k over slow secure channel.

Have G on both sides to get key of sufficient length to allow practical usage of one-time pad.



Even if every seed is expanded to be unique, hard cap.

world of pseudorandom seed
world of full length key.

ratio is fine

to make the claim of indistinguishability is a tall ask.

e.g.: a distinguisher can ask for $2^n + 1$ samples. $P(\text{collision}) = 1$

Distinguishers definitely exist, but polynomial time unknown.

this is not polynomial time

(Idea: Any test created by any polynomial time distinguisher will be passed by both.)

Proof: $\text{PRGs} \Rightarrow \text{one way function}$.

$$G(s) = y \quad (\text{intuitive implication})$$

$$\text{for some inputs, } G(s) = y \quad s \in P$$

now given y , find s st $G(s) = y$

|| hard as at least 2^n possibilities need to be tried

Proof: One way function \Rightarrow PRG.

\Rightarrow (G is one way)
 1. forward easy
 2. reverse hard

$$p = \text{prime}$$

$$\{1, 2, \dots, p-1\} \quad g \text{ is generator of } \mathbb{Z}_p^* \text{ if } \\ \mathbb{Z}^* = \{g^0, g^1, g^2, \dots, g^{p-1}\}$$

Discrete log problem

Given the group \mathbb{Z}_p^* and generator g , $y = g^n \bmod p$, find n .

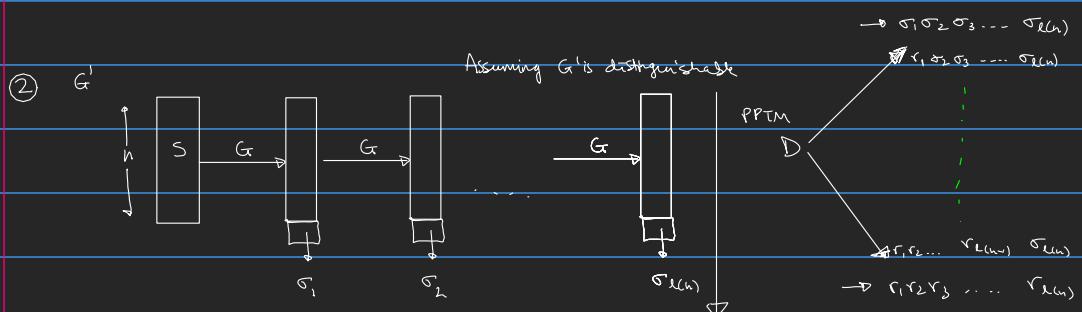
No efficient algorithm.

Can be used
as a PRG

$$\begin{array}{c|c} f: \{0,1\}^n \longrightarrow \{0,1\}^n & G: \{0,1\}^n \longrightarrow \{0,1\}^{(n+1)} \\ \downarrow \text{PRG } G & \downarrow \text{DLP } \Rightarrow G \\ \end{array}$$

① $\Rightarrow G': \{0,1\}^n \rightarrow \{0,1\}^{(n+1)}$

- Proof that
1. If I have length-preserving one-way function, I can make a PRG G that can expand its seed by 1 bit.
 2. If I have PRG G that can expand its seed by 1 bit, I can have PRG G' that can expand its seed by an arbitrary length.



if D can distinguish 1st from last,

$\exists i$ st i^{th} row and $(i+1)^{\text{th}}$ row is distinguished.

If G is PRG.

- It can distinguish that Δ (1st bit changed)
- $\Rightarrow (G')!$ was not pseudorandom (contradiction)

$\Rightarrow D$ has power to distinguish σ_i, r_i some

$\Rightarrow D$ can distinguish between σ_i, r_i one of the G 's.

but None can be , contradiction.

\therefore if G is PRG, G' is also PRG.

PROOF that if one-way functions exist, single bit expanding PRGs exist.

$$\textcircled{1} \quad G(s) = f(s) \parallel h(s)$$

length-preserving
PRG
trivial enough with
one-way function

looking at first n -bits, no polynomial function should be able to predict next n -bit with $\leq \frac{1}{2} + \text{negl}(n)$ probability.

given s , easy to compute $h(s)$

$$h(s) \leftarrow \begin{array}{l} \text{given } f(s), P[A(f(s)) = h(s)] \leq \frac{1}{2} + \text{negl}(n) \\ \text{guessing this is hard} \end{array}$$

$$h: \{0,1\}^n \rightarrow \{0,1\}$$

hard-core predicates of $f(\cdot)$

now, looking directly at DLP:

DLP_M: hardcore predicate.

$$f(x) = g^x \pmod{p}$$

$$\text{then } h(x) = \text{MSB}(x)$$

DLP: given $g^x \pmod{p}$

what is x

DLP(MSB): given $g^x \pmod{p}$
what is MSB(x)
 $[x \geq \frac{p-1}{2}]$

polynomial
reducing DLP_M to DLP

Now, we will show that if there is polynomial time algorithm for DLP(MSB) will also lead to a " " DLP, which is already discounted.

\textcircled{1} DLP_L: Given $g^x \pmod{p}$, what is the LSB(x)?

DPL_L $\in P$

\textcircled{2} DLP_L + DLP_M \Rightarrow DLP.

$$\text{Let } y = g^x \pmod{p}$$

Fermat's Little Theorem implies that

$$y^{p-1} \equiv 1 \pmod{p}$$

F.L.T.
 $a^{p-1} \equiv 1 \pmod{p}$
 if $\text{GCD}(a,p) = 1$

$$\therefore (g^x)^{p-1} \pmod{p} \geq 1$$

$$\text{Find } y^{\frac{p-1}{2}} \pmod{p} = g^{x \frac{(p-1)}{2}} \pmod{p}$$

x is even

x is odd

↓

$y^{\frac{p-1}{2}} \pmod{p} \equiv -1$

we know that $g^{\frac{p-1}{2}} \pmod{p} \equiv 1$

But $g^{\frac{p-1}{2}} \pmod{p} \equiv +1 \text{ or } -1$

now we know it cannot be 1,
it has to be -1.

$$g^x \bmod p$$

$x \text{ is even} \quad \quad x \text{ is odd}$

$\rightarrow g^{x/2} \bmod p \quad \quad g^{x+1} \bmod p.$

(SQRT) $\rightarrow g^{x/2} \bmod p, g^{\frac{x}{2} + \frac{p-1}{2}} \bmod p.$ } \rightarrow this won't tell me which was the original x unless I know p

\hookrightarrow gives 2 outputs. $\rightarrow p \equiv 3 \pmod{4}$ $\boxed{\text{if } x > p-1/2}$ which was our problem for DLP

can sqrt b.
done in poly
true?

W. yes, but 2
answers

Given x find $\sum_{z \in \mathbb{F}_p} z^2 \pmod{p} = x$.

$$\sum_{z \in \mathbb{F}_p} z^2 \equiv x^{\frac{p+1}{2}} \pmod{p}$$

$$z^2 \equiv x^{\frac{p-1}{2}} \pmod{p}$$

$$\Rightarrow x^{\frac{p-1+2}{2}} \pmod{p}$$

$$= \cancel{x^{\frac{p-1}{2}}} \cdot x \pmod{p}$$

because $x \in \mathbb{F}_p \quad x^2 = x \Rightarrow x^{\frac{p-1}{2}} \pmod{p}$

$$= 1 \pmod{p}$$

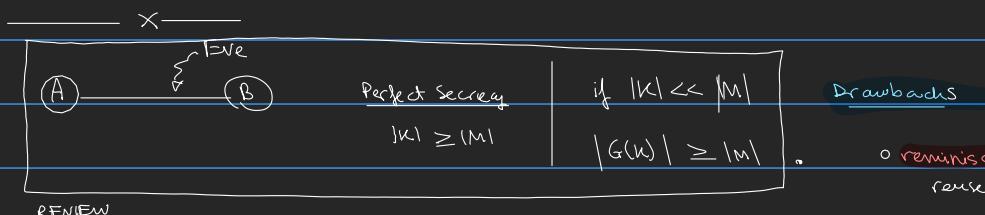
So if $p = 7$
 $\sqrt{2} = 2^{\frac{7+1}{4}} = 4$

28/1.26

Today

- o CPA-Security (prove that no deterministic encryption scheme can be CPA-secure)
[chosen Plaintext]

- o Quiz Paper



o reminiscent of one-time-pad: Cannot reuse $G(K)$

- needs state-awareness of what was passed before.

o Even if state is maintained,
it is very fragile

- messages sent in parallel will break the system
- if a packet is dropped, F

o will not work at internet scale
- the internet is stateless

Pseudorandom functions (PRF)

$G_r(k)$

random Access
to $G(k)$

required for
high scalability

Assume, atm, using this, we can have good, stateless encryption schemes.

((— it is not secure to run algorithm on text.))

CPA

$C \rightarrow m$ ciphertext only attack

if we also know

$c_1 \rightarrow m_1$

$c_2 \rightarrow m_2$

:

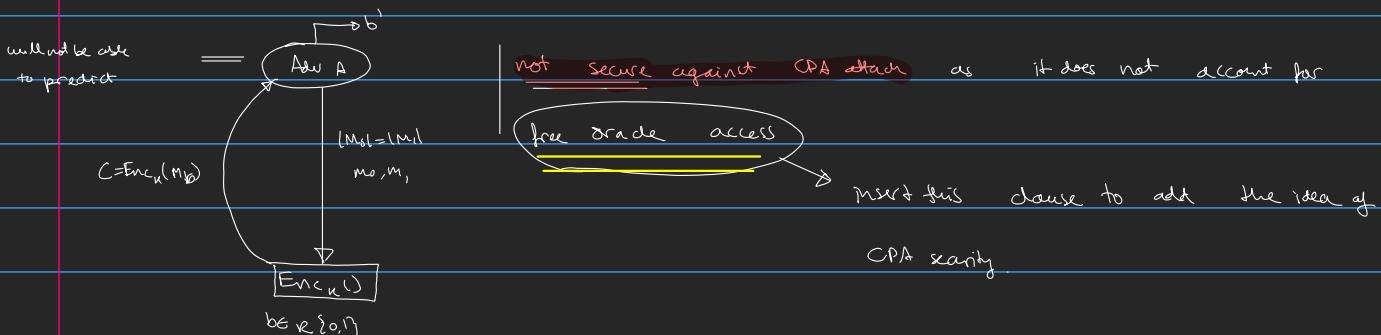
$c_t \rightarrow m_t$

||
adversary's choice

Q Why is such an attack practical?

— due to free availability of encryption servers, adversary can generate any number of ciphertexts for corresponding messages.

RECOLLECTING DEFN



Theorem: No deterministic encryption algorithm is CPA secure.

this will not affect

Stateful algo

Ex:

if $c = c_0, b = 0$

$(m_0, m_1) \rightarrow (c_0, c_1)$

$(m_0, m_1) \rightarrow (c_0, c_2)$

elif $c \neq c_0, b = 1$

Stateless is the only viable solution, but for blocking CPA attacks we need probabilistic encryption.

Q How to probabilistic encryption and deterministic decryption?

— trick: if $\text{Enc}_k = \{0,1\}^n \rightarrow \{0,1\}^n$ is deterministic.

Choose $r \in \{0,1\}^n$

$C = \langle r, m \oplus \text{Enc}_k(r) \rangle$

$\text{Dec}_k(r, v) = v \oplus \text{Enc}_k(r)$

((((DO NOT encrypt message directly)))

this is length doubling

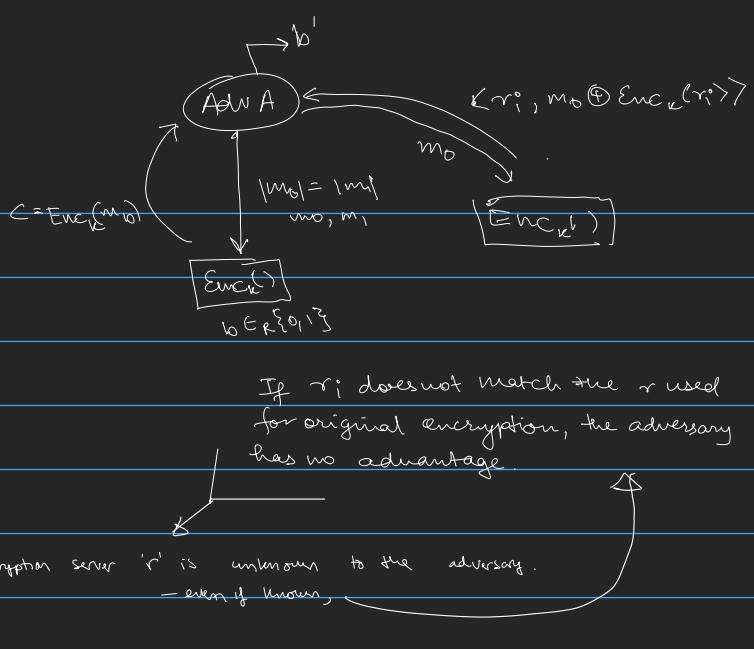
Modes of Operation

(Block Cipher)

① Cipher Block Chaining (CBC)

② output feedback mode (OFB)

③ Randomized Counter mode



OFB:

$m_1, m_2, m_3, \dots, m_t$
 $\langle r_1, c_1 \rangle, \langle r_2, c_2 \rangle, \dots, \langle r_t, c_t \rangle$

Define $r_i = \text{Enc}_K(r_{i-1})$

now, we give $\langle r_0, c_1, \langle r_1, c_2, \dots, c_t \rangle \rangle$ ALMOST LENGTH PRESERVING

converts one-wayness to a PKE.

Still in order of seconds. (t encryptions)

Randomized Counter Mode

$r_i = r_0 + i$ } is secure, will not go into prog now.

— allows for out-of-order decryption (good for connectionless networks)

CBC

— useful in data integrity

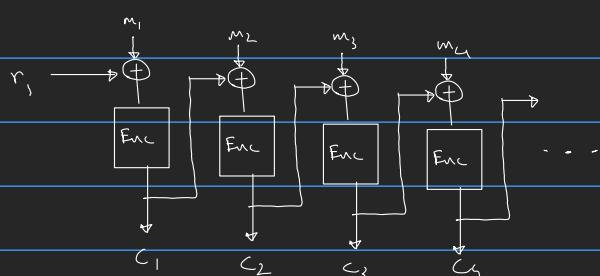
— pretty useless in encryption, compared to other 2.

if we have $m_1, m_2, m_3, \dots, m_t$

$\langle r=c_0, c_1, c_2, \dots, c_t \rangle$

$c_i = \text{Enc}_K(m_i \oplus c_{i-1})$

} has single r instead of several



Everything is enjoyable if it is unpredictable.

enjoyment or unpredictability

Quiz Qs

1. b) Show that shift cipher / vigenere cipher is perfectly secure if one alphabet/message length \leq key length

2. b) if $p-1 = s^{2^r}$, s is odd

then $(r+1)^{th}$ LSB is a hardcore predicate for DLP

Design a PRG using this

1. Shift cipher where one character is encrypted:

$$P(M=m | C=c) = P(k=k_0)$$

$$Enc_k(x) = (x+k) \% 26$$

for perfect secrecy,

$$P(M=m) = P(M=m | C=c)$$

$$P(M=m) = \frac{P(C=c | M=m) \cdot P(M=m)}{P(C=c)}$$

Now,

$$\text{form } P(C=c | M=m_0)$$

$$P(C=c) = P(C=c | M=m)$$

$$= P(k=k_0)$$

$$= \frac{1}{26}$$

(assume for M ,

$$P(C=c | M=m_0) = P(C=c | M=m_1)$$

(perfect secrecy)

$$\begin{aligned} & \text{for any } m=m_0, & & \text{if } m \in M \\ & P(C=c) = P(C=c | M=m_0) & & P(C=c) = \sum P(C=c | M=m_i) \cdot P(M=m_i) \\ & \text{otherwise for any } m=m_1, & & P(C=c) = \dots \end{aligned}$$

$$\Downarrow$$

$$P(C=c | M=m_0) = P(C=c | M=m_1)$$

2. b) if $p-1 = s^{2^r}$, s is odd

then $(r+1)^{th}$ LSB is a hardcore predicate for DLP

Design a PRG using this

- DLP = x given y, g, p , \mathbb{Z}_p^\times , $y = g^x \pmod{p}$
is one way function

$$G(f(u), h(u)) = f(u) \parallel h(u)$$

$$G(u) = u \parallel$$

$$G(u) = u+1 \Rightarrow G(u) \Rightarrow h(u)$$

Today

31.1.20

- * What is $F_k(r)$
 - Pseudo-random Functions
 - * PRFs exist iff PRGs exist
- $\text{OWF} \Leftrightarrow \text{PRG}$
 $\Leftrightarrow \text{PRF}$
 $\Leftrightarrow \text{CPA-secure}$
- not our concern now
MAC

Truly Random
 $\Rightarrow R_k(r)$

} Intuitively, first r returns random,
Subsequent r s give
Same answer for same y

A truly random function:

$$R_k : \{0,1\}^n \rightarrow \{0,1\}^{2^n}$$

$$(2^n)^{2^n} = 2^{n \cdot 2^n}$$

function.

↓
pick one at random
and choose that

If key is n bits [Keylength required for indexing R_k ?]

$n \cdot 2^n$ bits

$2^{n \cdot 2^n}$ functions

So again if you cannot distinguish pseudo random from truly random, then $F_k(r)$ can be pseudo random

definition of PRF, as a distinguisher from RF

Def of PRF

A function $F_k : \{0,1\}^n \rightarrow \{0,1\}^n$ is said to be a PRF if

(a) Efficiency: \exists an efficient deterministic algorithm that computes $F_k(x)$.

(b) Pseudo-randomness: \forall PPTM D , sufficiently large n :

$$\left| P[D(F)=1] - P[D(R)=1] \right| \leq \text{negl}(n)$$

$$\text{or } \left| P[D^{F_k(\cdot)}(1^n) = 1] - P[D^{R(\cdot)}(1^n) = 1] \right| \leq \text{negl}(n).$$

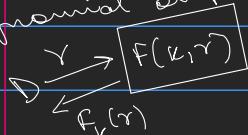
Theorem: PRFs exist iff PRGs exist

Step 1

$\text{PRF} \Rightarrow \text{PRG}$ (OFB mode of operation)

$$s_0 = s; s_i = F_k(s+i)$$

You can sample only polynomial outputs.



$$\text{PRG}_1 = G_1(s) = F_k(s_0) \parallel F_k(s_1) \parallel F_k(s_2) \parallel \dots$$

function \Rightarrow generator is trivial

$$F_k : \{0,1\}^n \rightarrow \{0,1\}^n$$

Step 2

$\text{PRG} \Rightarrow \text{PRF}$

Suppose we have $\text{PRG} \rightarrow G$, we use it to build PRF

$$G : \{0,1\}^n \rightarrow \{0,1\}^{2n}$$

$$\begin{aligned} G_0(s) &= \text{left half of } G(s) \\ G_1(s) &= \text{right half of } G(s) \\ G(s) &= G_0(s) \parallel G_1(s) \end{aligned}$$

Consider a function F_K taking 1-bit input and n bit output.

$$F_K(0) = G_0(K) \quad F_K(1) = G_1(K)$$

These are PRFs.

Now, for n -bit input

↓

$$F_K(r) = G_{r_{n-1}} \left(\dots G_{r_2} \left(G_{r_1} \left(G_{r_0}(K) \right) \right) \dots \right)$$

where,

$$r = r_0 r_1 r_2 \dots r_{n-1}, r_i \in \{0,1\}$$

$$K \in \{0,1\}^n$$

Construction: $\rightarrow G(K)$

$$G_0(K) \quad G_1(K)$$

$$\begin{matrix} G_{00}(K) & G_{01}(K) & G_{10}(K) & G_{11}(K) \\ \checkmark & \checkmark & \checkmark & \checkmark \end{matrix}$$

⋮

2^n

Proof of PRFness

If there is a distinguisher which distinguishes $G_{01}(K)$ from truly random, then it can distinguish $G_{00}(K)$ which is not possible as $G_{00}(K)$ is a pseudo random generator.

base assumption

Consider OWF $\rightarrow \text{DLF}$ $f(n) = g^n \bmod p$

take N of DLF ($\text{MSB}(n)$)

Consider $S_i = S$

$$S_i = f(S_{i-1}) = g^{S_{i-1}} \bmod p$$

Construct a PRG

$$G(S) = h(S_1) \parallel h(S_2) \parallel h(S_3) \parallel \dots$$

↓

$$G(S) = \text{MSB}(S_1) \parallel \text{MSB}(S_2) \parallel \text{MSB}(S_3) \parallel \dots$$

Construct a PRF

$$F_K(r) = G_{r_{n-1}} \left[\dots G_{r_0}(K) \dots \right]$$

↓

Construct a

$$\text{Enc}_K(m) = \langle r, F_K(r) \oplus m \rangle \rightarrow \text{CPA secure encryption}$$

$$\text{Enc}_K(m_1 \dots m_t) = \langle r_0, F_K(r_0 + i) \oplus m \rangle \leftarrow \text{randomised counter.}$$

$$\text{AES}_K(r) = c$$

$$|K| = 128 \text{ bits}$$

$$|C| = |r| = 128 \text{ bits}$$

can be 256 bits

P = NP

Strong AI

Algorithmic

PRG

OWF

OWF

Hash

Sig-Benc

CPA

CCA

MAC

Sig

ZKP

Trapdoor OWF

PKC

2-party

ppIR

sMPC

Crypto

Math

Worst-case hardness

not on average

Avg-case hardness

OWF not exist

Min Crypt

PKCs do not exist

Signatures exist

Heuristic

Pessiland

04/02/2020

Review

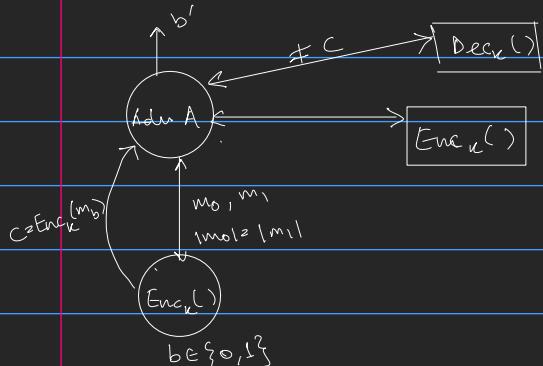
- CPA - secure Enc
- PRF

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

Today

- # CCA - security
- # MAC

chosen-Ciphertext Attack



CCA - sec

$$\begin{aligned} & \forall \text{ PPTM A} \\ & P[b' = b] \leq \text{negl}(n) \end{aligned}$$

Consider

$$\begin{aligned} m_0 &= 0^n \\ m_1 &= 1 \oplus 0^{n-1} \end{aligned}$$

$$c_0 = \langle r, m_0 \oplus F_k(r) \rangle$$

$$c_1 = \langle r, m_1 \oplus F_k(r) \rangle$$

$$c' = \langle r, \text{Toggle MSB } (m_0 \oplus F_k(r)) \rangle$$

$$\begin{aligned} \text{Dec}_k(c') &\rightarrow \begin{cases} 0 & \text{if } \text{Dec}_k(c') = m_1 \\ 1 & \text{if } \text{Dec}_k(c') = m_0 \end{cases} \end{aligned}$$

$$S_0, b' = \begin{cases} 0 & \text{if } \text{Dec}_k(c') = m_1 \\ 1 & \text{if } \text{Dec}_k(c') = m_0 \end{cases}$$

$$P[b = b'] = 1$$

This system of encryption is malleable (with a known change to c , the change in m can be predicted).

CCA - security (To be done later)

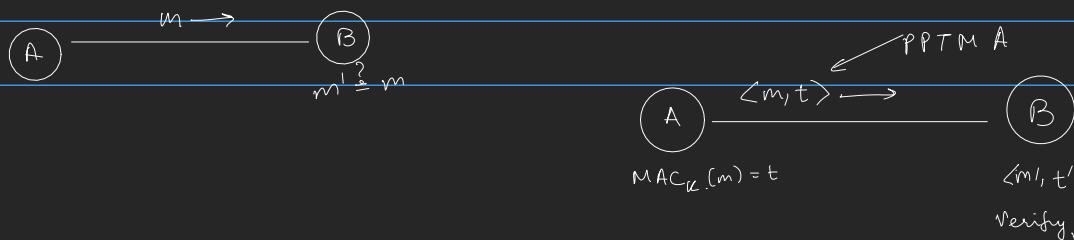
Avalance effect: explore this

MAC - Message authentication codes.

$$\boxed{\text{CPA} + \text{MAC} \Rightarrow \text{CCA}}$$

So what is MAC?

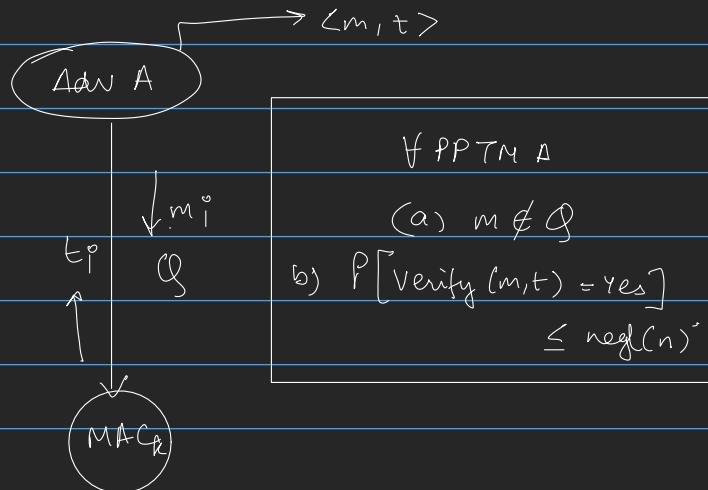
MAC (Message Authentication Codes)



$$\text{Verify}_k(m', t') = \text{Yes/No}$$

Security of MAC

- Solves data integrity



Encrypt - Then - Authenticate

$$\text{Ciphertext} = \langle c, t \rangle$$

$c \xrightarrow{\text{CPA-Enc}_{K_1}(\cdot)}$ $t \xrightarrow{\text{MAC}_{K_2}(\cdot)}$

Dec. = check for 'No-change'

$$\text{Verify}_{K_2}(c, t) = \text{Yes}$$

if No - 1

$$\text{Yes } \text{Dec}_{K_1}(c) = m.$$

Designing MACs from PRFs

If the only messages to be authenticated

$$F_K : \{0,1\}^n \rightarrow \{0,1\}^n$$

are $\{0,1\}^n$

↳ then straight forward

1 way: $t = F_K(m)$ people don't do this // usually $|m| \gg |key|$

Some approaches that failed miserably:

1. $t = F_K(\oplus m_i)$, where $m = m_1, m_2, \dots, m_n$ X FAIL

- easy to find other m' to match

2. $t = \langle t_1, t_2, \dots, t_n \rangle$ where $t_i = F_K(m_i)$ X FAIL

- can drop packets (drop $\langle m_n, t_n \rangle$) or permute

3. $t_i = F_K(i||m_i)$ X FAIL

- can drop packets at the end

4. $t_i = F_K(i||m_i)$ X FAIL

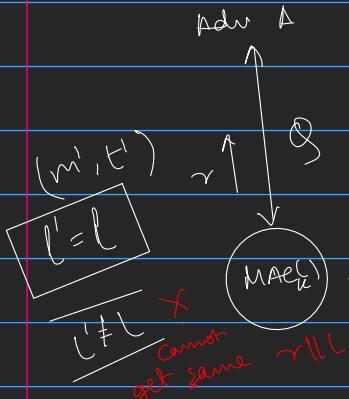
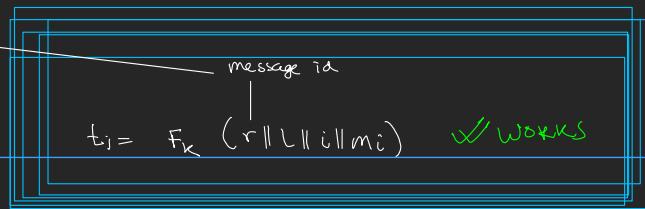
- $i = p_1, p_2, \dots, p_n \rightarrow t_1, t_2, \dots, t_n \left\{ \begin{array}{l} m_1: p_1, q_2, p_3, q_4, \dots \\ m_2: p_2, q_1, p_3, q_3, \dots \\ \vdots \\ m_n: p_n, q_1, p_2, q_2, \dots \end{array} \right. \}$

Exponentially many messages possible.

Send seq. no.

Send length

message ID is
randomly gen



$$m = m_1 m_2 \dots m_l, (t_1, t_2, \dots, t_l)$$

$$t'_i = F_K(r || l || i || m'_i)$$

$$m = m_1 m_2 \dots m_l$$

$$m' = m'_1 m'_2 \dots m'_l$$

$$\begin{array}{c} m \neq m' \\ \exists i \quad m'_i \neq m_i \end{array}$$

But this is not practical as the length of tag is as long as the message itself.

CBC MAC_K

\rightarrow cipherblock chaining.

$$m_1, m_2, \dots, m_l$$

$$t_i = F_K(t_{i-1} \oplus m_i)$$

$$t = t_a$$

$$t_0, t_1, t_2$$

$$(t_1, c_1), (t_2, c_2)$$

$$\textcircled{1} \quad t = F_{K_2}(t_a)$$

$$t_a = \text{CBCMAC}_{K_1}(m)$$

✓ used for stream of data where l is not known beforehand

$$\textcircled{2} \quad K' = F_K(l) \quad t = t_a \quad \text{CBCMAC}_{K'}(m)$$

$$\textcircled{3} \quad \text{prepend the length}$$

$$m' = l || m_1 || \dots || m_l$$

$$t = \text{CBC MAC}(m')$$

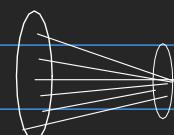
✓ In practice

11.2.20

Collision Resistant Hashing

$$h: \{0,1\}^* \rightarrow \{0,1\}^n$$

A collision: $x \neq y, h(x) = h(y)$



by definition, there are lot of collisions.

Idea: design a function that has collisions, but not to a PPTM adversary.

Security param.: you have a family of hash functions, indexed.

$$h^s$$

FORMAL
DEFⁿ of
(family f) hash function.

(family f) hash function.

→ PPTM A,

$$P[A^n = \langle x,y \rangle \mid n \neq y, h^s(n) = h^s(y)] \leq \text{negl}(s)$$

Uses

- Password hash
- Message digest
- Hash chain

if DLP hard, construction for hash function is correct.

first thing noticed: no matter how hard you try, you cannot get past generic Birthday Attack.

choose an n that is at least double the security needed, as GBA can halve due to square root.

generic Birthday Attack.

• Set of N elements

• Pick q at random (with replacement)

$\Rightarrow Q \quad P(\text{picking an element twice or more})$

$$\text{Then } \frac{q(q-1)}{4N} \leq P[\text{Collision}] \leq \frac{q^2}{2N}$$

{ if $q > \sqrt{N}$... are in trouble }

$$q = \sqrt{N} \approx 2^{n/2}$$

$$N = 2^n$$

Prob {Union bound?} scrapped
 $P \leq \frac{\binom{q}{2}}{N}$

$$P[\text{Collision}] = P[N_{C_1}] \cdot P[N_{C_2} \mid N_{C_1}] \cdot P[N_{C_3} \mid N_{C_1, C_2}] \cdots P[N_{C_{q-1}} \mid N_{C_{q-2}}]$$

$$= \prod_{i=1}^{q-1} \left(1 - \frac{i}{N}\right)$$

$$= \prod_{i=0}^{q-1} \left(1 - \frac{i}{N}\right)$$

$$P(\text{Collision}) = 1 - \prod_{i=0}^{q-1} \left(1 - \frac{i}{N}\right)$$

$$\text{if } 0 < x < 1 \quad 1-x \leq e^{-x} \leq 1-\frac{x}{2}$$

Now,

$$\prod_{i=1}^{q-1} \left(1 - \frac{i}{N}\right) \leq \prod_{i=1}^{q-1} e^{-\frac{i}{N}}$$

$$\leq e^{-\sum_{i=1}^{q-1} \frac{i}{N}}$$

$$\leq e^{-\frac{q(q-1)}{2N}}$$

$$\leq 1 - \frac{q(q-1)}{4N}$$

$$\therefore P(\text{coll}) \geq \frac{q(q-1)}{4N}$$

Merkle-Damgard Transform

given a hash function with some compression ratio, build hash function with any compression ratio.

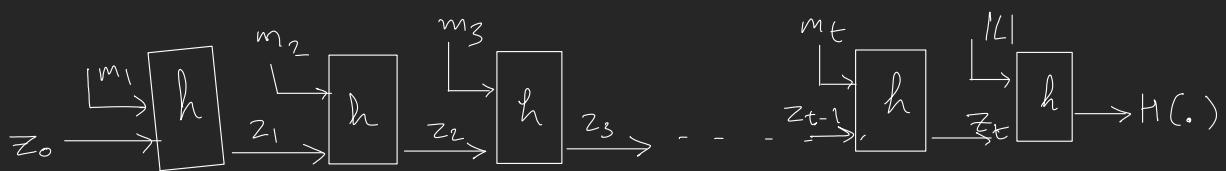
$$h : \{0,1\}^{2^n} \rightarrow \{0,1\}^n$$

If h is collision resistant, then

$$H : \{0,1\}^* \rightarrow \{0,1\}^n$$

Merkle-Damgard will give us collision resistant H .

$$H(m_1 | m_2 | m_3 | \dots | m_t)$$



$$H(m_1 \dots m_t) = H(m'_1 m'_2 \dots m'_{t'})$$

If $t \neq t'$: then $h(z_t | t) = h(z'_t | t')$

Since $t \neq t'$ then $z_t | t \neq z'_t | t'$

If $t = t'$: $h(z_t | t) = h(z'_t | t)$

If $z_t \neq z'_t$ then, $h(z_t | t) = h(z'_t | t)$
and $z_t | t \neq z'_t | t$ } collision

If $z_t = z'_t$:

If $m_t \neq m'_{t-1}$: $h(z_{t-1} | m_t) = h(z'_{t-1} | m'_t)$
and $m_t \neq m'_t$

If $m_t = m'_{t-1}$:

needs doing this till the end.

Finally we find $m_1 | m_2 | \dots | m_{t-1} | m_t = m'_1 | m'_2 | \dots | m'_{t-1} | m'_t$
which are different at i^*

Now

$$h(z_{i^*} | m_{i^*}) = h(z'_{i^*} | m'_{i^*})$$

but $z'_{i^*-1} | m'_{i^*} \neq z'_{i^*} | m'_{i^*}$

we have
found
collision
then.

Designing $h : \{0,1\}^{2n} \rightarrow \{0,1\}^n$.

$$z = g^k \bmod p$$

Define $h(x, y) = g^x \cdot z^y \bmod p$.

g is generator of \mathbb{Z}_p^*

$h(x_1, y_1) = h(x_2, y_2)$ but $x_1, y_1 \neq x_2, y_2$

$$y_1 \neq y_2$$

$$g^{x_1} \cdot z^{y_1} \equiv g^{x_2} \cdot z^{y_2}$$

$$g^{x_1 - x_2} \equiv z^{y_2 - y_1}$$

$$g^{x_1 - x_2} \equiv g^{k(y_2 - y_1)}$$

$$(g^{x_1 - x_2})^{(y_2 - y_1)} \equiv (g^{k(y_2 - y_1)})^{(y_2 - y_1)}$$

$$g^{\frac{x_1 - x_2}{y_2 - y_1}} \equiv g^k$$

$$k = \frac{x_1 - x_2}{y_2 - y_1}$$

but this means that
you have found
discrete log of z
So h is collision
resistant function

Public Key Cryptography

Review of Symmetric Key Crypto

① Two Relaxations to Shannon

Neg. P(adv)
PPTM Adv

② Pseudo-Random Generator (PRG)

DLP

③ CPA Security

No det. sec is CPA secure.

④ Probabilistic Encryption

⑤ Pseudo-Random Function

for prob enc

⑥ Modes of Operation

making it practical

⑦ PRF iff PRG

⑧ PRF, PRG from DLP, Hard-core predicate

block ciphers (Strong Pseudorandom Permutation (PRP))
↓
PRF bijection

⑨ CCA security

CPA not enough

⑩ MACs

CCA can be done using concept of MAC

⑪ CBC MAC

⑫ CPA + MAC → CCA

⑬ Hashing, Collision Resistance

⑭ Merkle-Damgård Transform

Convert low compression coll. res hash to arbitrary.

⑮ Hashing, from DLP

use DLP to get low compression coll. res hash

PRF

how to make invertible without breaking pseudorandom property

use "Feistel structure"

if you have

$f: \{0,1\}^n \rightarrow \{0,1\}^n$ (PRF)

Data Encryption Standard (DES)

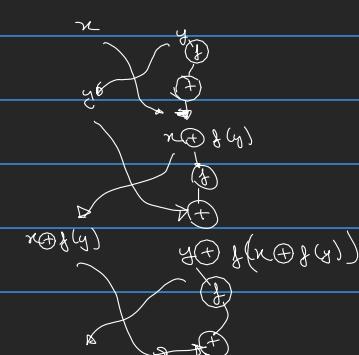
proven

4 rounds of Feistel Structure, with

without provable PRF, cannot say.

a non-invertible PPF, then we get

if we have 2 inputs, (x, y)



a bijected PRF(PRF) = block cipher

_____ x _____ x _____ x _____ x

Public Key Crypto



No Secure Channel

[No key shared a prior]

Tuesday:
Diffie-Hellman

Coin theory: need:

• for inter-router communication, a system would need to store n^2 unique keys.

• with PKC: a single key can be used.

Diffie-Hellman on Key Establishment

Relaxations: Same as Shannon, - PPTM Adv
- Negl error

Assumed Statement: Computational Decision Diffie-Hellman assumption

DDH Assumption

Group (cyclic)

\mathbb{Z}_p^* , g

Decision DDH Assumption

given g^a, g^b, g^{ab} world 1 } if not distinguishable,
 g^a, g^b, r world 2 } DDH holds

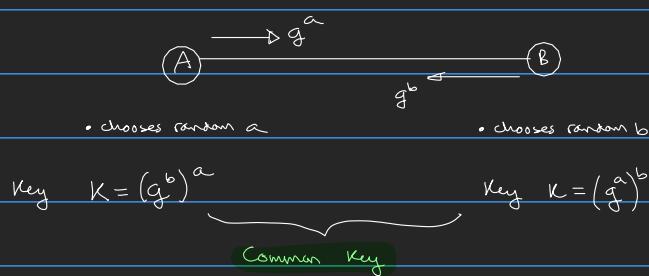
Computational DDH Assumption

Statement

$$\forall \text{PPTM } A, \quad |P[A(g^a, g^b, g^{ab}) = 1] - P[A(g^a, g^b, r) = 1]| \leq \text{negl}(\log p)$$

proof leads to vacuous result: DH is secure because it is assumed to be secure.

Code for secure key exchange:



g^x : operate g x times

Now, to show this is secure

Can be proven under DDH assumption: if g^a, g^b, g^{ab} cannot be found.

Secure under assumption that
it is secure

(Why is this Turing Award worthy?)

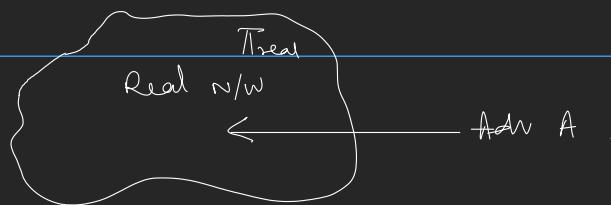
because this led to explosion of crypto protocols.

lots of things can be true if DDH assumption is true. so...

"Anything that can be done insecurely can be done securely if DDH is true"

Universal function: trapdoor one-way function

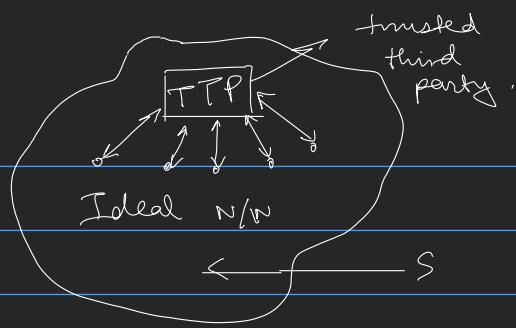
what is SIM definition (different from indistinguishable defn)



Aside: "you're students, you're supposed to read left, right, center. But times have changed. or maybe not.. when I was a student, even I didn't read much"

Ideal N/W is secure.

Real N/W is not.



There is a TTP.

Whatever you want to do,
you let TTP do the work and
it then returns the answer.

Treat is a secure protocol for
"whatever" if for PPTM A in the
real network,
 $\exists S \in \text{Ideal N/W}$
network s.t. $\stackrel{c}{\equiv} \text{View}(A) = \text{View}(S)$

For every real life adversary, there will exist a corresponding
adversary in the ideal life such that their views are identical.
strict equality \Rightarrow Shannon's definition.

18.02.2020

Midsem Review:

(A) Breaking Historical Ciphers

- (A.1) Shift Cipher
- (A.2) Mono-alphabetic Substitution cipher
- (A.3) Vigenère Cipher
- (A.4) Kerckhoff's Principle.

(B) Perfect Secrecy

- (B.1) Defining PS
- (B.2) Proof that Vernam Cipher is PS
- (B.3) Limitations of PS ($|K| \geq |M|$)

(C) Relaxations to PS

- (C.1) PPTM Adversary
- (C.2) Negligible, but non-zero, error probability.
- (C.3) Precise Assumptions (1-way functions exist)

(D) Security against ciphertext-only Attacks

- (D.1) Defining Pseudo-random Generators (PRG)
- (D.2) PRG \Leftrightarrow One Way functions (Designing PRG using DLO)
- (D.3) Enc. $C = G(k) \oplus m$ works.

(E) CPA security

- (E.1) Defining CPA-secure enc.
- (E.2) No det. scheme is CPA-sec
- (E.3) Probabilistic Enc.
- (E.4) Pseudo-random functions
- (E.5) PRF \Leftrightarrow PRG
- (E.6) $C = \langle r, M \oplus F_k(r) \rangle$ works
- (E.7) Modes of operation

(F) CCA-sec and MACs

- (F.1) Defining CCA-sec enc and MACs
- (F.2) MACs from PRF
- (F.3) CS-CMACs
- (F.4) CCA-sec Enc from CPA-sec enc.
and secure MAC.

(G) Hashing

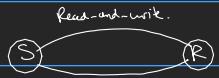
- (G.1) Defining Collision Resistance
- (G.2) Birthday Attacks
- (G.3) Merkle-Damgård Transform
- (G.4) Designing Hash Functions from DLP

(H) Key Exchange

- (H.1) Diffie-Hellmann Protocol
- (H.2) DDH - assumption

Questions for Mid

H:



Adv has access to both channels, and write access to one.

Challenge: You do not know which channel is read only for adv and which

based on:

DDH Assumption
R&W access to channel by adv.

IS read-unit.

PPTM

Design a protocol st. S, R have keys, but Adv does not, even with
tw access to any channel of their choice.

Rewriting with back ground

in standard scheme, there is only one channel, read only. [Diffie-Hellman Works]

but it fails if adversary has write access.

Once key is shared, even write access can be dealt with.

in this question, there are two channels. One is read+write for the adversary, one is read-only. we do not know which is which. [PPTM]

Design a protocol st. S, R have keys, but Adv does not, even with
tw access to any channel of their choice.

G.

$$h: \{0,1\}^{2n} \rightarrow \{0,1\}^n \xrightarrow{\text{MDT}} H: \{0,1\}^* \rightarrow \{0,1\}^n$$

Can be used with arbitrary compression ratios.

start with an h which is, say, $h: \{0,1\}^{3n} \rightarrow \{0,1\}^{2n}$

there should still be an MDT st $H: \{0,1\}^* \rightarrow \{0,1\}^{2n}$

big aside on

"what is theory"

TODD PROOF

F. FORMALIZE

Break the Base-CBCMAC scheme.

↳ just take message, do CBCMAC, take answer

Replay Attacks are allowed

Q: Why? ↗

G. OWF \Leftrightarrow PRG \Leftrightarrow PRF \Leftrightarrow MAC \Leftrightarrow hashing \Leftrightarrow block cipher.

Q: Start somewhere, build everything else.

Solutions

1. Encryption scheme : $\langle \text{Gen}, \text{Enc}, \text{Dec}, \mathcal{M} \rangle$

Shift Cipher

$$\text{Gen} := k \leftarrow [0, 25]$$

$$\text{Enc} := c_i = (m_i + k) \% 26$$

$$\text{Dec} := m_i = (c_i - k) \% 26$$

$$\mathcal{M} = [a, z]^*$$

Substitution cipher

$$\text{Gen} := k \leftarrow \text{Permutation of } a \dots z$$

$$\text{Enc} := c_i = k(m_i)$$

$$\text{Dec} := m_i = k^{-1}(c_i)$$

$$\mathcal{M} = [a, z]^*$$

$$k: a \rightarrow c$$

$$b \rightarrow n$$

Vigenere cipher (key length l)

$$\text{Gen} := k \leftarrow [0, 25]^l$$

$$\text{Enc} := c_{kn+i} = (m_{kn+i} + k_i) \% 26 \quad | \quad n \in \mathbb{W}$$

$$\text{Dec} := m_{kn+i} = (c_{kn+i} - k_i) \% 26 \quad | \quad n \in \mathbb{W}$$

$$\mathcal{M} = [a, z]^*$$

2. let an alphabet permutation be $p: [a, z] \rightarrow [a, z]$

$$\text{Key: } p_1, p_2, \dots, p_t$$

1. If we know key length:

- separate cipher text into t buckets:

$$c_1, c_{t+1}, c_{2t+1}, \dots$$

$$c_2, c_{t+2}, c_{2t+2}$$

- For each bucket, use frequency analysis attack

- count all character frequencies.

- sort characters by frequencies

- sort alphabet by known real-world frequencies

} line up, form the key

2. Finding key length:

3. Perfect secrecy:

$$\frac{P(M=m \mid C=c)}{P(C=c \mid M=m)} = P(M=m)$$

$$\downarrow \quad \begin{array}{c} P(C=c \mid M=m) P(M=m) - P(M=m) \\ P(C=c) \end{array}$$

$$P(M=m) = 0.5$$

$$P(M=m) = 0.1$$

$$P(C=c \mid M=m) = P(C=c)$$

$$P(C=c) = \sum_{m \in M} P(C=c \mid M=m) \cdot P(M=m)$$

$$P(C=c \mid M=m') = P(C=c \mid M=m)$$

$$P(M=m \mid C=c) = P(M=m)$$

$$P(M=m' \mid C=c) = P(M=m') \quad \text{QED}$$

refuted.

4. a) $f(u) < 1$ $g(u) < 1$

$$\text{pol}_1(u)$$

$$\text{pol}_2(u)$$

$$f(u) \cdot g(u) < 1$$

$$\text{pol}_1(u) \cdot \text{pol}_2(u)$$

QED

b) $f(u) = \frac{1}{2^n}$ $g(u) = \frac{1}{4^n}$ ← counter example

$$f(u)/g(u) = \frac{2^{2n}}{2^n} = 2^n \neq \frac{1}{\text{pol}(u)}$$

QED

5. $P[A(f(x)) = y \mid f(x) = f(y)] \leq \text{negl}$

$$P[A(g(x)) = y \mid g(x) = g(y)] \leq \text{negl}$$

a) $P[A(f(x) \oplus g(x)) = y \mid f(x) \oplus g(x) = f(y) \oplus g(y)] \leq \text{negl}(x)$ ← wrong.

for any given $a, b, c \in \mathbb{Z}$ st $a \oplus b = c \oplus d$. and $|a| = |b| = |c| = |d|$

$$P(a \oplus b = c \oplus d) \approx \left(\frac{1}{2}\right)^{|x|}.$$

a). Counter example: $f = g$.

then $f \oplus g = 0$ for all inputs $x \Rightarrow$ not one way function

$$b) h(n) = f(f(n))$$

assume A for $h(n)$.

$f(n)$ is oWF } $f(n)$ is a bijection of the same space on
 $\{x, n \text{ have same space}\} \Rightarrow |f(n)| = |x|$ itself
 $\Rightarrow A \text{ for } h(n) \text{ is also for } f(n)$
[contradiction]

This f' is one-way. In fact, this holds even if only f is one-way (regardless of g , as long as g is efficiently-computable). To see this, fix a PPT adversary \mathcal{A}' and let

$$c) \epsilon(n) \stackrel{\text{def}}{=} \Pr[\mathcal{A}'(f'(x)) \text{ outputs an inverse of } f'(x)],$$

where the probability is taken over uniform choice of x and the random coins of \mathcal{A}' . Consider the following PPT adversary \mathcal{A} : given input y_1 (which is equal to $f(x_1)$ for randomly-chosen x_1), choose random x_2 , compute $y_2 := g(x_2)$, and run $\mathcal{A}'(y_1 \| y_2)$. Then output the first half of the string output by \mathcal{A}' . It is not hard to see that (1) the input $y_1 \| y_2$ given to \mathcal{A}' is distributed identically to $f'(x_1 \| x_2)$ for randomly-chosen x_1, x_2 . This implies that \mathcal{A}' inverts its input with probability $\epsilon(n)$. Furthermore, (2) whenever \mathcal{A}' successfully inverts its input, \mathcal{A} successfully inverts its own input. We conclude that \mathcal{A} outputs an inverse of y_1 with probability at least $\epsilon(n)$, showing that ϵ must be negligible.

$$d) h(n_1, n_2) = (f(n_1), n_2)$$

$$A, \quad \Pr[A(h(n_1, n_2)) = y_1, y_2 \mid h(y_1, y_2) = h(n_1, n_2)] \not\leq \text{negl}.$$

$$y_2 = n_2$$

$$\boxed{f(y_1) = f(n_1)} \text{ not poss by PDTMA}$$

= contradiction.

6.

G is a function, s.t $\text{length of output} = l(\text{length of input})$??

a) s is an n -bit binary string (random)

$$y_0 = G(s) \quad y_1 \leftarrow_R \{0, 1\}^{l(n)}$$

b) b is a random bit ($b \in \{0, 1\}$)

c) give y_{sb} to A . A outputs b' (if y_0 , output 0. if y_1 , output 1)

$$\Leftrightarrow \text{constructing } A: A(y_{sb}) = D_s(y_b) \begin{cases} 0 & \text{if } y_b = y_0 \\ 1 & \text{if } y_b = y_1 \end{cases}$$

For any adversary A interacting with the given experiment, we have that

$$\begin{aligned} \Pr[b' = b] &= \Pr[b' = 0 \mid b = 0] \cdot \Pr[b = 0] + \Pr[b' = 1 \mid b = 1] \cdot \Pr[b = 1] \\ &= \frac{1}{2} \cdot \Pr[A(G(s)) = 0] + \frac{1}{2} \cdot \Pr[A(r) = 1] \\ &= \frac{1}{2} \cdot (1 - \Pr[A(G(s)) = 1]) + \frac{1}{2} \cdot \Pr[A(r) = 1] \\ &= \frac{1}{2} + \frac{1}{2} \cdot (\Pr[A(r) = 1] - \Pr[A(G(s)) = 1]). \end{aligned}$$

So $|\Pr[b' = b] - \frac{1}{2}| \leq \text{negl}(n)$ iff $|\Pr[A(r) = 1] - \Pr[A(G(s)) = 1]| \leq \text{negl}(n)$.

$$\Pr[b' = b] \leq \frac{1}{2} + \text{negl}(n) \quad \Pr(A(y_0) = 0) + \Pr(A(y_1) = 1) \leq 1 + \text{negl}(n)$$

$$\Pr(A(y_0) = b) \leq \frac{1}{2} + \text{negl}(n) \quad \text{And } \Pr(A(y_0) = 0) = 1 - \Pr(A(y_0) = 1) \quad \text{since } G \text{ is PRA},$$

$$\Pr(A(y_0) = 0) \leq \frac{1}{2} + \text{negl}(n) \quad \Pr(A(y_0) = 1) \leq \frac{1}{2} + \text{negl}(n)$$

$$\Pr(A(y_1) = 1) \leq \frac{1}{2} + \text{negl}(n) \quad \leq \frac{1}{2} + \frac{1}{2} \cdot \left[\Pr(A(y_0) = 1) - \Pr(A(y_0) = 0) \right]$$

Up leap

$$\Pr(D_s(y_b) = 1) \leq \frac{1}{2} + \text{negl}(n)$$

$$\Pr(D_s(y_0) = 1) \leq \frac{1}{2} + \text{negl}(n)$$

7. 1 year: $1 \cdot 21 \times 10^{34} \text{ J}$

flip 1 bit: $5 \cdot 8 \times 10^{-23} \text{ J}$

flip 256 bits: $2^{256} \times 5 \cdot 8 \times 10^{-23} \text{ J}$

time in years: $\frac{2^{256} \times 5 \cdot 8 \times 10^{-23} \text{ J}}{1 \cdot 21 \times 10^{34} \text{ J}} = 2^{256} \times 4.79 \times 10^{-57} \text{ years}$

$2^{10} > 10^3$
$2^{140} > 10^{57}$
$2^{140} \approx 1.6 \times 10^{57}$

$$2^{256} \approx 3 \times 2^{140}$$

$$3 \times 2^{66}$$

Not Possible

8. Let G be a function that maps strings of length n to strings of length $2n$. Define

$$\gamma(n) \stackrel{\text{def}}{=} \Pr[\text{the } (n+1)^{\text{th}} \text{ bit of } G(x) \text{ is equal to } 1]$$

where the probability is taken over random choice of $x \in \{0,1\}^n$. Prove that if G is a pseudorandom generator, then there is a negligible function ϵ with $\gamma(n) \leq 1/2 + \epsilon(n)$. (Give a formal proof, not just an intuitive argument.)

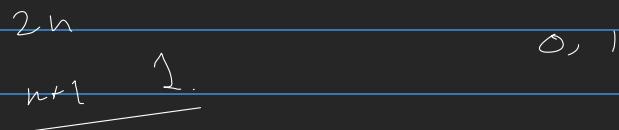
8. G is a PRG.

in a purely random $2n$ -bit string, distribution of 0,1 bits is 1:1

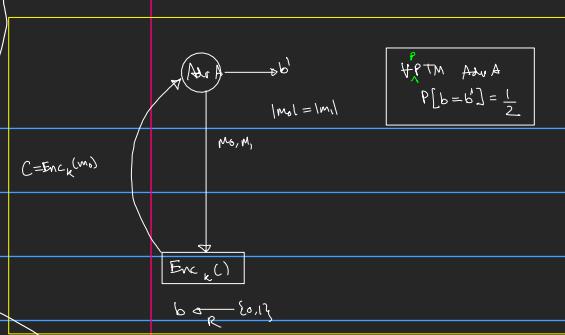
in G , if $\gamma(n) \geq \frac{1}{2} + \text{negl}$, then $\gamma(n)$ can act as a distinguisher
for G (differing distribution compared to random)

contradicts the assumption that G is a PRG.

$$\therefore \gamma(n) \leq \frac{1}{2} + \text{negl}.$$



9.



Equivalent Definition of Perfect Secrecy.

An encryption scheme is **cipher-text-only secure** if
 $\forall \text{PTM } A, P[b = b'] \leq \frac{1}{2} + \text{negl}(n)$
 (P of selecting correct message from the message space is negligibly more than random)

$$q. a) m_1 = c_1 \oplus g(k)$$

$$m_2 = c_2 \oplus c_1$$

$$b) \text{ eq. def: } P[b = b'] \leq \frac{1}{2} + \text{negl}.$$

$$\text{but, } m_0 = m_a m_b \quad m_1 = m_c m_d \quad b = m_0 \text{ or } m_1$$

$$\text{we get } c' = c_n \oplus y$$

we can get m_y from $c_n \oplus y$

$\hookrightarrow m_y = m_b \text{ or } m_d \Rightarrow \text{we know the message}$

$$\Rightarrow P[b = b'] = 1 \quad \text{Violation}$$

QED.

10. a) OWP \rightarrow PRF.

$$G^{n+1} = \text{OWP} \parallel h(\text{OWP})$$

$$G^{n+1} \rightarrow G^*$$



$p_1 p_2 \dots p_n$

$v_1 v_2 \dots v_n$

A
 $A v_i p_i$

Contra

$$b) 2^{l(n)} \rightarrow 2^n$$

c) Consider a function F_K taking l bit input and n bit output.

$$F_K(x) = G_0(x) \quad F_K(x) = G_1(x)$$

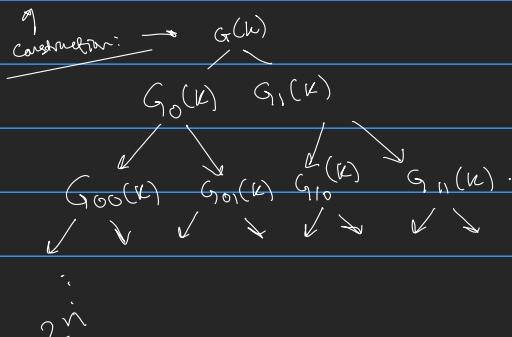
These are PRFs.

Now, for n -bit input

$$F_K(r) = G_{r_{n-1}}(\dots G_{r_2}(G_{r_1}(G_{r_0}(K))) \dots)$$

Constructed PRF

where,
 $r = r_0 r_1 r_2 \dots r_{n-1}, r_i \in \{0, 1\}$



Proof of PRFness

If there is a distinguisher which distinguishes $G_{r_0}(K)$ from truly random, even it can distinguish $G_0(K)$ which is not possible as $G_0(K)$ is a pseudo random generator.

ii) PRGs \Rightarrow PRFs

$$G: \{0,1\}^n \rightarrow \{0,1\}^{2n}$$

Let $G_0(s) = \text{left half of } G(s)$

$G_1(s) = \text{right half of } G(s)$

Hence $G(s) = G_0(s) || G_1(s)$

$$k \in \{0,1\}^n, r \in \{0,1\}^n$$

$$F_k(r) = G_{r_{n-1}} [G_{r_{n-2}} [\dots G_{r_2} [G_{r_1} [G_{r_0}(k)]] \dots]]$$

$$\text{TP: } \forall \text{ PPTM } D / p(D(F_k(r)) = 1) - p(D(F_{k'}(r)) = 1) \leq \text{negl}(n)$$

Assume to the contrary that \exists a PPTM D .

Base Case: For $|r| = 1$, $F_k(r) = G_{r_0}(k)$

Clearly if D can distinguish $F_k(r)$ then it can distinguish G .

Inductive Hypo: We cannot distinguish upto r of length n .

Consider r of length $n+1$.

If we can distinguish $F_k(r)$

PRF

a). how to make invertible without breaking pseudorandom property

use "Feistel structure"

if you have $f: \{0,1\}^n \rightarrow \{0,1\}^n$ (PRF)

Data Encryption Standard (DES)

without provable PRF, cannot say.

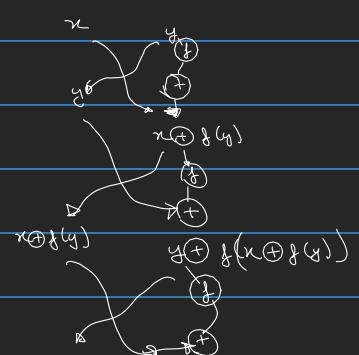
But 16 rounds of Feistel might do.

problem

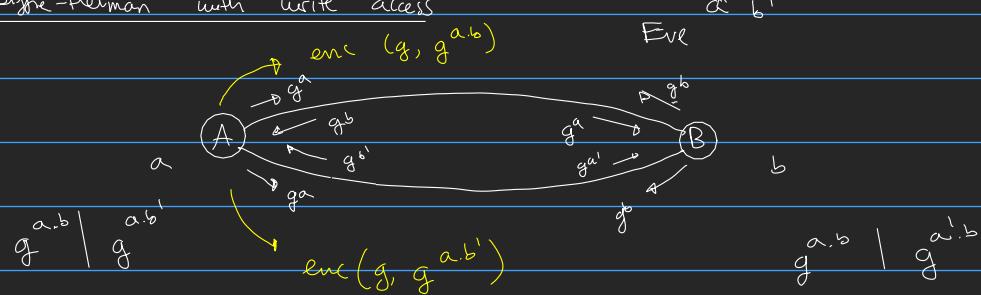
4 rounds of Feistel Structure, with
a non-invertible PRF, then we get

a bijective PRF (PRP) = block cipher

if we have 2 inputs, (x, y)



Diffie-Hellman with write access



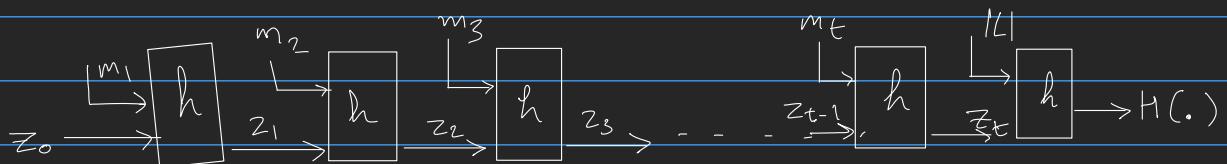
$$g^{ab} \curvearrowright g^{ab}$$

$$g^{ab} \times g^{a'b}$$

$$g^{ab} \times g^{ab}$$

$$g^{ab} \times g^{a'b}$$

$$2. |m_i| = n \quad |z_i| = 2n$$



3.

$$(m, t)$$

$$(m', t)$$

$$m' [0] \oplus t \parallel m$$

attack

