Intro to Modern Cryptography
– J. KATZ + Y. LINDELL

POIS

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

Will be shown with standard examples.

eg1: Hashing password.
theoretically impossible to be perfect if length is not infinite.

eg2: secure communication
@ t_0 info(R) = info(eve)
@ $t > t_0$ info_rec(R) = info_rec(eve)

eg3: Data integrity
If m was sent, modified to m' → It is the same to receiver.
and if m' " " ,
└ cannot thus identify.

Eg of non infosec → infosec

- problem in distributed computing ∴ now an Infosec problem.

Solution: use signatures. ⇒ 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.

FUNDAMENTAL
- Information security is God
- all nontrivial works of science must include info sec

Random Words
– Hamming Distance

–logical nogo

Sources of Impossibility.
① Computational Hardness [Resource Complexity]
② Practical Uncertainties (Only buggy code are secure.)
③ Natural Limits (Speed of light, Quantum Stuff)
④ Logical / Philosophical Impossibilities

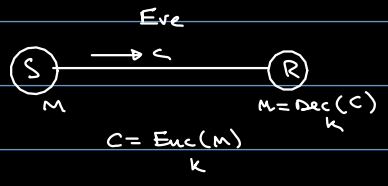
7.1.20

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

$C = (x + 3) \bmod n_c$

M = message
C = coded message

n_c = no. of characters in alphabet.

Brig 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 Reasoning

1. Algorithms are reverse engineerable.

   eg: $h(x) = 5x + c$

   Attacker can feed $x_1, x_2, \ldots x_n$ and see that for all outputs $h(x_i)\ldots$

   $8 \mid h(x_i) - h(x_j)$. And then solve for c.

2. Updation / Recovery Complexity.

   if password revealed in secure system: change pass.
   if algo " in obscurity " : !! fixed.

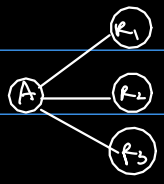
3. Secure Memory is costly.

   Proof in modern times: UTM is algo and just input is key.

   — bad information storage efficiency.

ASK ATHREYA

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

- → bug exists
- → nonethical people exist
- → only " " search for bug
- → " " " will find "
- → big 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.

→ If keyspace is smol, attack ez. (brute force)

first principle learnt:
Principle of large key space

ATTACK

1. Can be broken by humans
2. Autobreaking:
   - frequency analysis

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

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

Now compute $\sum_{i=0}^{25} (p_i q_{i+k})$ — wrong guess: $1/26$; correct guess: $\simeq 0.065$

Next iteration:

Monoalphabetic Substitution Cipher

- Diff alphabets shift by different amts.
- no repetitions allowed
- basically, permutation

for brute force: 26! keys to search ← applying first principle

ATTACK

$\forall i \ \exists j : q_j \simeq p_i$ ⇒ 1. Sort q_is, 2. Sort p_is — since distribution is same, eg: $p_a = q_n$, $p_e = q_t$, $p_t = q_b$

Issue: susceptible to frequency attacks.

Next iteration

Vigenere Cipher

1. does not maintain frequency

eg: hello
sease
zilh_

FAILED ATTACKS
- brute force: too many keys
- freq. anal: freq. not maintained.

Can be broken if:
1. key length known
2. " " is findable.

ATTACK

PART 1:

if we know length of key, eg: 3 | C_0 C_3 C_6 ....

partition ciphertext into (length) parts | C_1 C_4 C_7 ....

then shift cipher attack on all (length) parts. | C_2 C_5 C_8 ....

PART 2:

Guess an l.

take a string $C_0 C_l C_{2l} \ldots$

check if $\sum_{i=0}^{25} q_i^2 = \sum_{i=0}^{25} p_i^2$

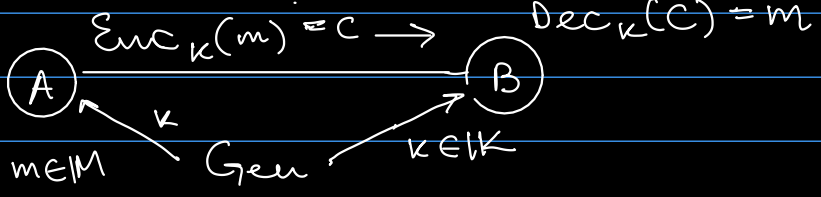
easy.

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
$\langle Gen, Enc, Dec, M \rangle$



$Dec_K(Enc_K(m)) = m$

Perfectly secret encryption scheme.

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

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

Schemes that meet this are largely impractical

Giving definitions is easier than proving, but workable definitions, not so much.

We will prove one time pad is perfectly secret.

$$P[M=m \mid C=c] = P[M=m]$$

Random Variables.

Vernam Cipher

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

Gen : $k \xleftarrow{R} \{0,1\}^n$

Enc : $c = m \oplus k$

Dec : $m = c \oplus k$.

- correctness is fairly obvious and not shown here.

$a \oplus a = 0$

$m \oplus a \oplus a = m \oplus 0 = m$

$\text{Dec}_k(\text{Enc}_k(m))$

$= m \oplus k \oplus k = m$

(Proved)

PROOF

First showing that definition of perfect security is equivalent to

- $\forall$ p dist over $\mathbb{M}$

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

$P[C=c \mid M=m] = P[C=c]$

have to do an iff.

Suppose this holds.

$$P[C=c \mid M=m] = P[C=c]$$

multiply both sides.

$$\left\{ \frac{P[M=m]\ P[C=c \mid M=m]}{P[C=c]} \right\} = P[M=m]$$

Bayes'

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

Second

showing $\forall c = \mathbb{C}, \forall m_0, m_1 \in M^2$ is eq to (1)

$P[C=c \mid M=m_1] = P[C=c \mid M=m]$

2 → 1 is trivial ($P(C=c \mid M=m) = P(C=c)$, LHS=RHS)

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

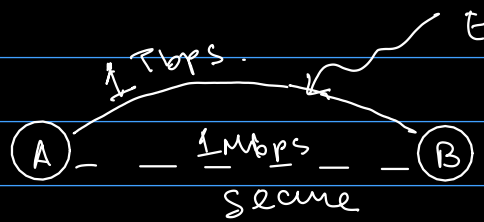
$$= p \sum_{m \in M} P[M=m] = p$$

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:

$$\text{LHS} = P[C = c \mid M = m_0]$$
$$= P[C = m_0 \oplus K] = P[K = C \oplus m_0]$$
$$= \frac{1}{2^n} = \text{RHS} \quad (P \text{ is not dependent on } m_0)$$



Uses of Vernam Cipher:

i) To save your shady business from the feds, encrypt the data and decrypt 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.

Slow secure channel + fast insecure channel $\Rightarrow$ Fast secure channel

Impossible if perfect security.

Two impossibilities

Public Key cryptography

Only insecure channel $\Rightarrow$ slow secure channel

Both are complementary.

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

Mini version 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 " " " " key.

Missing here: M is compressible.

Shannon did: $H(|K|) \geq H(|M|)$

We use a hacky bypass to this.

ASK ATHREYA FOR INTUITION.

Proof :

Suppose not.

$$|K| < |M|$$

(the contrary)

we will show this directly implies this cannot be perfectly secret.

some ciphertext c

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

now, $|D| \leq |K| \ \ \therefore < |M|$

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

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

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

but we said $P(M = m^*) \neq 0$

$\Rightarrow$ Scheme is not perfectly secret

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

$\therefore$ 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
- associative
- identity
- inverse

for group G,

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

Cyclic group if $a \in G$, and $G = \{ a^0, a^1, a^2, \ldots \}$ eg. of $(\mathbb{Z}_n, \cdot_n)$ groups

Next: Ring, Integral Domain, Field.

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

a) $(R, +)$ is a commutative group — if $a.b = b.a \ \forall a, b \in R$

b) Closure : $a.b \in R \ \forall \ a, b \in R$

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

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

Ring need not be commutative with $\cdot$

Commutative Ring is also " " $\cdot$

eg. of Ring: $(\mathbb{Z}, +, \cdot)$ is a ring.

$(\mathbb{Q}, +, \cdot)$ " " "

Zero Divisors: For a ring R, a, b s.t $a(\neq 0) \in R$, $\exists\ b \in R$, $b \neq 0$, $ab = 0$ or $ba = 0$

eg: $(\mathbb{Z}_n, +_n, \cdot_n)$

Integral Domain: A commutative ring with no zero divisors.

Division Ring: An integral domain s.t $(R - \{0\}, \cdot)$ is a group.

Field: $(F, +, \cdot)$

Finite field: Field where set F is finite.

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

b) $(F - \{0\}, \cdot)$ " " "

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

Not a field: $(\mathbb{Z}, +, \cdot)$, $(\mathbb{Z}_n, +_n, \cdot_n)$ generally

field: $(\mathbb{R}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$, $(\mathbb{Z}_n, +_n, \cdot_n)$ if $n = p$, $(\mathbb{Z}_n^*, +_n, \cdot_n)$

Characteristic of Integral Domain (I)

least positive integer m s.t $\underbrace{m \cdot a}_{\text{m times}} = 0 \quad \forall a \in I$

if such m does not exist, characteristic of I is 0.

Lemma: If $(F, +, \cdot)$ is finite, $\text{char } F = p$ (for some prime)

$\exists\ n$ $\cdot$ $\exists$ atleast 0, 1. now $1+1,\ 1+1+1,\ 1+1+1+1 \ldots$.

sums have to be repeated, as finite.

$i.1 = j.1$

$(i-j).1 = 0$

$n = i - j$

n is prime $\cdot$ if n is not prime, $n = ab$.

how do we know $a, b \in F$

$(a.b).1 = 0$

$\Rightarrow (a.1)(b.1) = 0$

$\Rightarrow a.b = 0$ $(a, b$, non-zero$)$

$\Rightarrow$ one of a, b zero divisor

contradictory to definition

$\Rightarrow n$ is prime.

Now, $(\mathbb{Z}_p, +_p, \cdot_p)$ is a field char $= p$

if for any given F, $\text{char } F = p$, $\mathbb{Z}_p (= \mathbb{Z}_{/p\mathbb{Z}}) \subset F$

- if $(F, +, \cdot)$ is finite, $\text{char } F = p$ (for some prime p)

Let $(F, +, \cdot)$ be a finite field $|F| = q$, let $F \subset K$ where K is also a finite field

then K has q^n elements where n is dimension of K over F

$\{v_1, v_2, \ldots v_n\}$ is a basis of K over F

$a \in K$, $a = a_1 v_1 + a_2 v_2 + \ldots a_n v_n$, $a_i \in F$

$|K| = q^n$

$(K, +, \cdot)$, $\text{char } K = p$

eg: $(\mathbb{C}, +, \cdot)$ is a vector space over $(\mathbb{R}, +, \cdot)$ $\quad |K| = p^n$

given a prime, field with exactly those many elements. Read Herstein

Elliptic Curves

Can be defined over finite fields

- gives an abelian group.
- multiple curves can be taken.

Should not have multiple roots.

— $y^2 = x^3 + ax + b$ defined over F

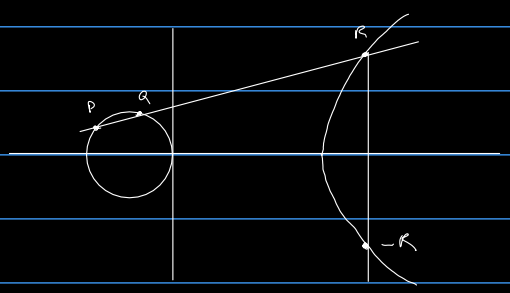
$P, a \quad l = \overline{PQ}$

l intersects this curve exactly at one more point

$P + Q = -R$

$P + O = O + P$ — O is point at infinity

$R - R = O$



How this helps in crypto:

2 parties agree on:

1. Field F
2. Eq. of elliptic curve
3. Random B

Alice — F, a, b, B — Bob

Alice: private mB → $m(nB)$

Bob: private nB → $n(mB)$

published

Shared key

21. 1. 20.

Review

1940s
⇒ Shannon's Perfect Secrecy and Pessimistic Result [P.S ⇒ |K| ≥ |M|]

1970s - early 80s
- Two famous Relaxations | One Universal Assumption.

Class lectures 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 error"] → 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

Used to define "negligible error"

Practical Adversaries: PPTM in Security Parameter

Negligible Function: A function μ(n) is said to be negligible in n, if it is smaller than the inverse of any polynomial for a sufficiently large n.

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

∴ Secrecy (Relaxed): ∀ PPTM Adversary; P [A can break π] ≤ negl()

What exactly do we need for reasonable negligibility — and why fixed fractions 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".

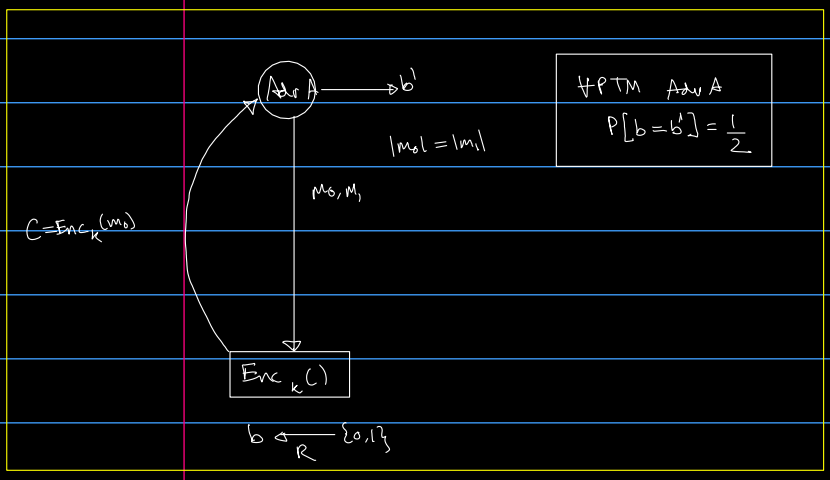
for examples exponential growth.

$\forall p(), \exists n_0$ 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

$\forall$ PTM A, $P(b = b') \leq \frac{1}{2} + negl(n)$

(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 need this definition.

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 object: One way functions.

Does this exist? No fun clue. There is a lacuna.

Fields, at this point, trifurcated.

① Heuristic Security.

(Secure because we have assumed it is)

eg: 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 → ϕ | taught when it exists

② Provable Security

we will give a conditional proof.

Definition of One-Way Function

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

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

b) Hard to invert : $\forall$ PPTM A

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

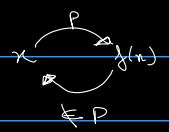
P(successfully inverting function) is negligible in previously defined terms.

Trio of impossibility beaters: Two famous relaxations + 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,

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

Verifier: given proof, we can verify it.

$L \in NP$ if L has an efficient verifier

if $x \in L$, $\exists$ at least 1 v which works. A verifier V is said to verify L if

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

Now,

if $\exists$ p(t) NDTM, $\exists$ efficient verifier ??? ASK ATREUS

Reverse process has efficient verifier

∴ it must be NP.