COS433/Math 473: Cryptography

Mark Zhandry
Princeton University
Spring 2017

Announcements

Reminder: Homework 1 due tomorrow 11:59pm

Submit through Blackboard

Homework 2 will hopefully be posted tonight

Last Time

Deterministic and Randomized Encryption

Perfect Secrecy = Perfect Semantic Security

OTP

Statistical Distance

Given two distributions D_1 , D_2 over a set X, define

$$\Delta(D_1,D_2) = \frac{1}{2}\sum_{x} | Pr[D_1=x] - Pr[D_2=x] |$$

Observations:

$$0 \le \Delta(D_1, D_2) \le 1$$

$$\Delta(D_1, D_2) = 0 \iff D_1 = D_2$$

$$\Delta(D_1, D_2) \le \Delta(D_1, D_3) + \Delta(D_3, D_2)$$

$$(\Delta \text{ is a metric})$$

Perfect Secrecy [Shannon'49]

Definition: A scheme (Enc,Dec) has perfect secrecy if, for any two messages $\mathbf{m_0}$, $\mathbf{m_1} \subseteq \mathbf{M}$

 $Enc(K, m_0) \stackrel{d}{=} Enc(K, m_1)$

Identical distributions

Random variable corresponding to uniform distribution over **K**

Random variable corresponding to encrypting $\mathbf{m_1}$ using a uniformly random key

Perfect Secrecy [Shannon'49]

Definition: A scheme (**Enc,Dec**) has **perfect** secrecy if, for any two messages $\mathbf{m_0}$, $\mathbf{m_1} \subseteq \mathbf{M}$

 Δ (Enc(K, m₀), Enc(K, m₁)) = 0

Another View of Statistical Distance

Theorem:
$$\Delta(D_1,D_2) \ge \epsilon \text{ iff } \exists A \text{ s.t.}$$

$$| Pr[A(D_1) = 1] - Pr[A(D_2) = 1] | \ge \epsilon/2$$

Terminology: for any A, $|Pr[A(D_1) = 1] - Pr[A(D_2) = 1]|$ is called the "advantage" of A in distinguishing D_1 and D_2

Another View of Statistical Distance

Theorem:
$$\Delta(D_1,D_2) \ge \varepsilon$$
 iff $\exists A$ s.t.
 $Pr[A(D_1) = 1] - Pr[A(D_2) = 1] \ge \varepsilon/2$

To lower bound Δ , just need to show adversary A with twice that advantage

Obtaining Perfect Secrecy: The One-Time Pad

Key space $K = \{0,1\}^n$ Message space $M = \{0,1\}^{\leq n}$ Ciphertext space $C = \{0,1\}^{\leq n}$

Enc(k, m) =
$$k_{[1, |m|]} \oplus m$$

Dec(k, c) = $k_{[1, |c|]} \oplus c$

Example:

k = 0011010110 m = 100101 c = 101000

Correctness:

Dec(k, Enc(k, m)) =
$$k\oplus(k\oplus m)$$

= $(k\oplus k)\oplus m$
= $0\oplus m$
= m

Limitations of OTP

It is only one-time

Try to encrypt two messages, security will fail

Enc(k,m₀)
$$\oplus$$
 Enc(k,m₁)
= (k \oplus m₀) \oplus (k \oplus m₁)
= m₀ \oplus m₁

Key length ≥ message length

 Limited use in practice: if I can securely transmit nbit key, why don't I just use that to transmit n-bit message?

Today

Multiple message security

Using the OTP more than once

- Stateful encryption
- Limitations

Multiple messages with stateless encryption

- Impossibility of perfect secrecy
- Security parameter
- Statistical secrecy

For today, assume both parties have extremely long shared secret key

When encrypting **m** s.t. **|m| << |k|**, don't have to throw away all of **k**

- Only k_[1, lml] has been used
- Use rest of k to encrypt next message

Syntax for Stateful Encryption

Syntax:

- Key space K, Message space M, Ciphertext space C
- State Space \$
- Init: $\{\} \rightarrow S$
- Enc: K×M×S → C×S
- Dec: K×C×S → M×S

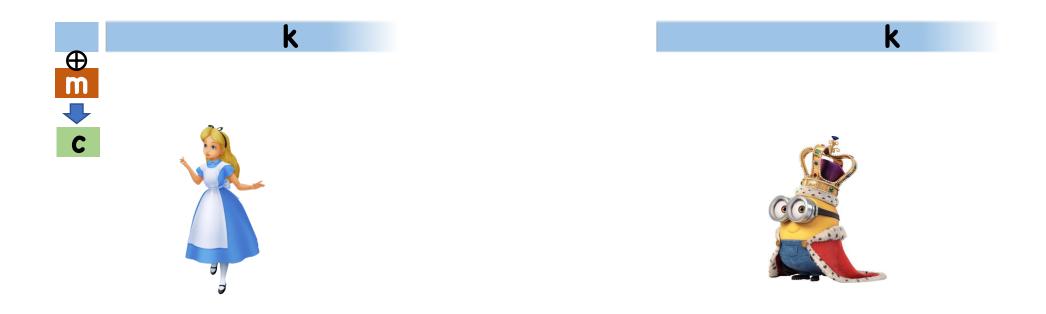
```
State<sub>0</sub> \leftarrow Init()
(c<sub>0</sub>, state<sub>1</sub>) \leftarrow Enc(k,m<sub>0</sub>,state<sub>0</sub>)
(c<sub>1</sub>, state<sub>2</sub>) \leftarrow Enc(k,m<sub>1</sub>,state<sub>1</sub>)
```

•••

m



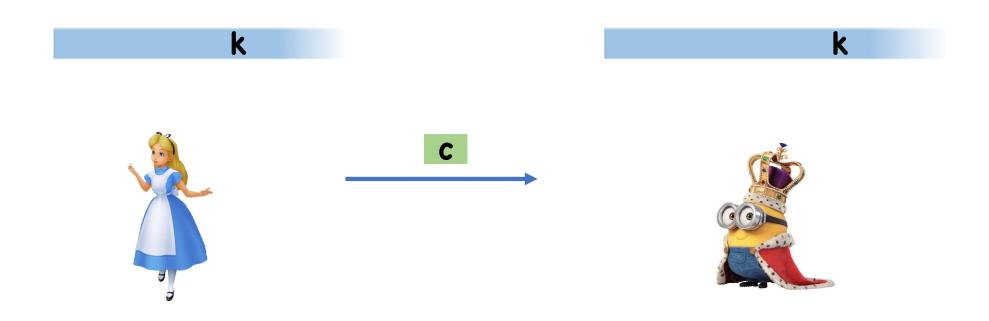


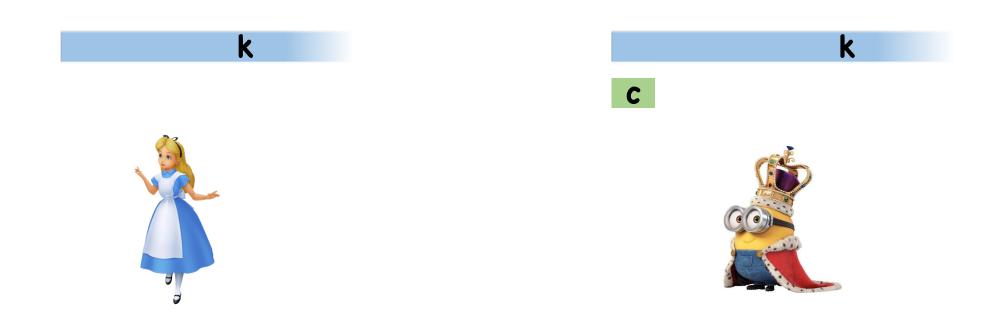


C

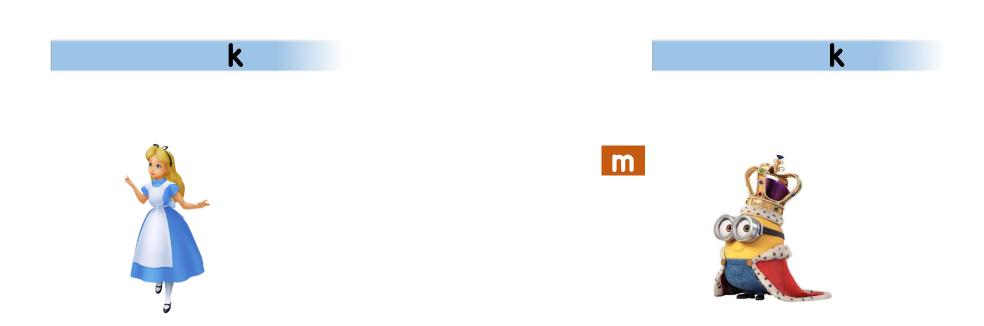


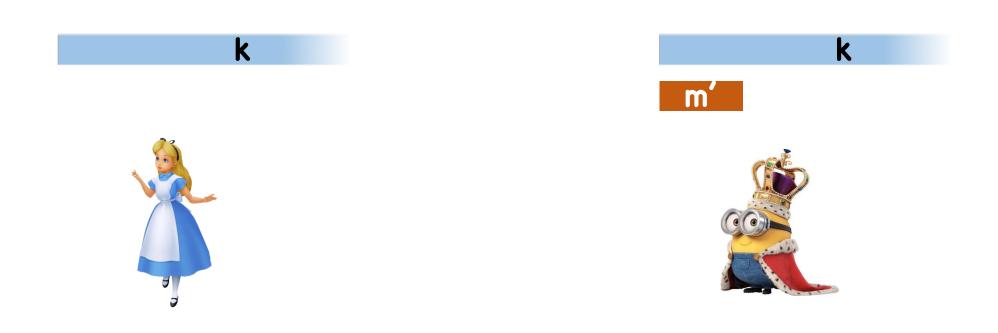


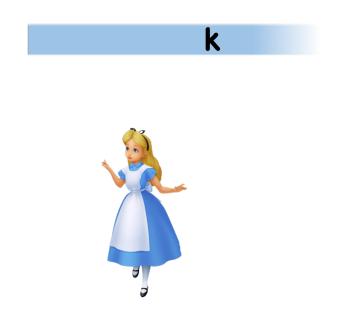


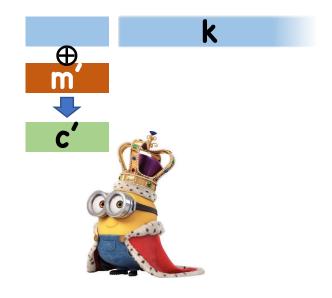


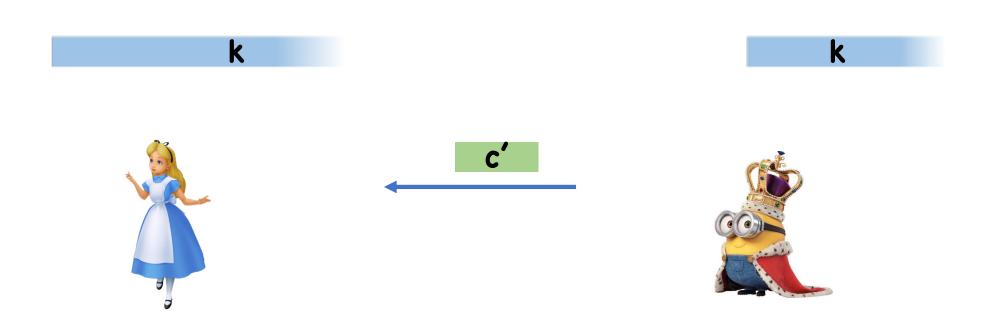


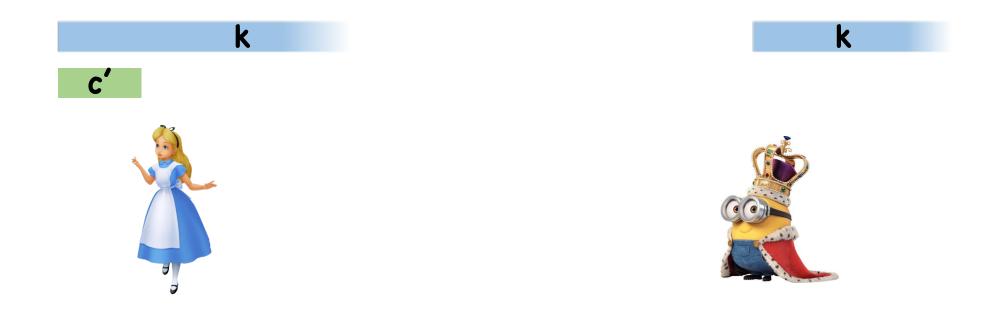


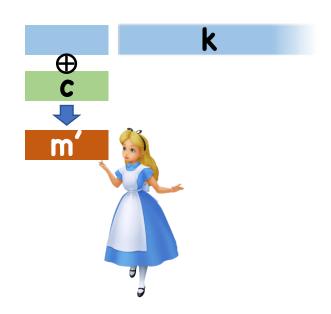












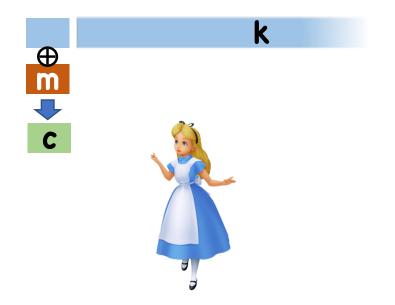


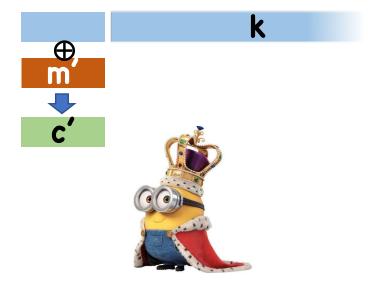


In real world, messages aren't always synchronous

What happens if Alice and Bob try to send message at the same time?

They will both use the same part of the key!





(

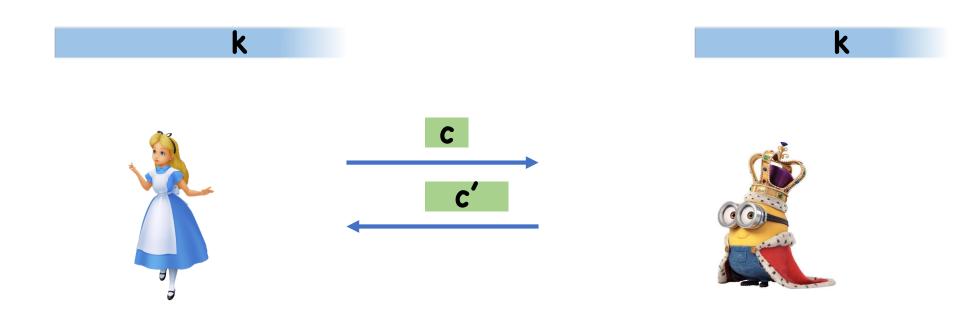
K

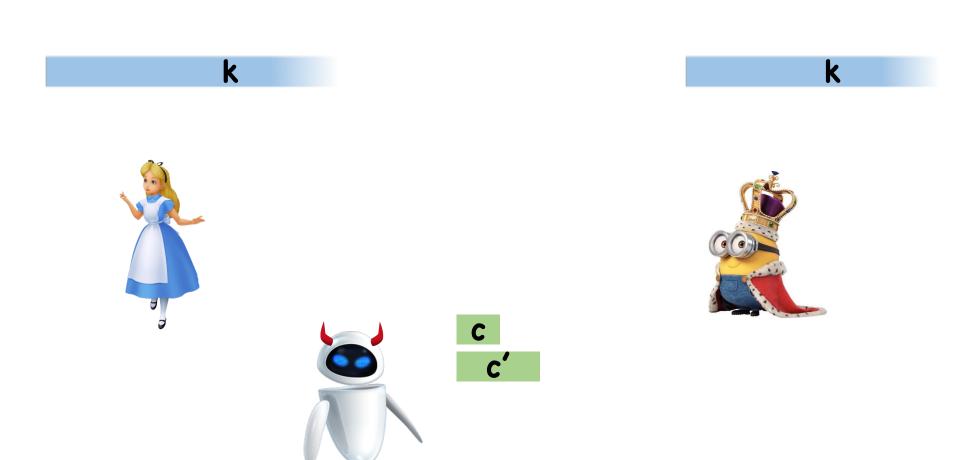
C



C'



















Solution

Alice and Bob have two keys

- One for communication from Alice to Bob
- One for communication from Bob to Alice

Can obtain two logical keys from one by splitting key in half

• Ex: odd bits form $k_{A\rightarrow B}$, even bits form $k_{B\rightarrow A}$

 $k_{A\rightarrow B}$

 $k_{B\rightarrow A}$

 $k_{A\rightarrow B}$

 $k_{B\rightarrow A}$





Still A Problem

In real world, messages aren't always synchronous

Also, sometimes messages arrive out of order or get dropped

Need to be very careful to make sure decryption succeeds

These difficulties exist in any stateful encryption

 For this course, we will generally consider only stateless encryption

Back to Stateless Encryption

Syntax:

- Key space K
- Message space M
- Ciphertext space C
- Enc: $K \times M \rightarrow C$
- Dec: K×C → M

Perfect Security for Multiple Messages

Definition: A stateless scheme (**Enc,Dec**) has **perfect** secrecy for **n** messages if, for any two sequences of messages $(m_0^{(i)})_{i \in [n]}$, $(m_1^{(i)})_{i \in [n]} \in M^n$

$$(Enc(K, m_0^{(i)}))_{i \in [n]} \stackrel{d}{=} (Enc(K, m_1^{(i)}))_{i \in [n]}$$

Notation: $(f(i))_{i \in [n]} = (f(1), f(2), ..., f(n))$

Ex:

$$M = C = \mathbb{Z}_p$$
 (p a prime)
 $K = \mathbb{Z}_p^* \times \mathbb{Z}_p$
 $Enc((a,b), m) = (am + b) \mod p$
 $Dec((a,b), c) = (c-b)/a \mod p$

Q: Is this perfectly secure for two messages?

Ex:

$$M = \mathbb{Z}_p$$
 (p a prime)
 $C = \mathbb{Z}_p^2$
 $K = \mathbb{Z}_p^2$
 $Enc((a,b), m) = (r, (ar+b) + m)$
 $Dec((a,b), (r,c)) = c - (ar+b)$

Q: Is this perfectly secure for two messages?

Theorem: No stateless encryption scheme* can be perfectly secure for two messages

Easier case:

Theorem: No stateless *deterministic* encryption scheme can be perfectly secure for two messages

Proof of Easy Case

Let (Enc, Dec) be stateless, deterministic

Let
$$\mathbf{m}_0^{(0)} = \mathbf{m}_0^{(1)}$$

Let $\mathbf{m}_1^{(0)} \neq \mathbf{m}_1^{(1)}$

Consider distributions of encryptions:

• (
$$c^{(0)}$$
 , $c^{(1)}$) = ($Enc(K, m_0^{(0)})$, $Enc(K, m_0^{(1)})$)
 $\Rightarrow c^{(0)} = c^{(1)}$ (by determinism)
• ($c^{(0)}$, $c^{(1)}$) = ($Enc(K, m_1^{(0)})$, $Enc(K, m_1^{(1)})$)
 $\Rightarrow c^{(0)} \neq c^{(1)}$ (by correctness)

Generalize to Randomized Encryption

Let (Enc, Dec) be stateless, deterministic

Let
$$\mathbf{m}_0^{(0)} = \mathbf{m}_0^{(1)}$$

Let $\mathbf{m}_1^{(0)} \neq \mathbf{m}_1^{(1)}$

Consider distributions of encryptions:

- (
$$c_{(0)}$$
 , $c_{(1)}$) = (Enc(K, $m_{0}_{(0)}$), Enc(K, $m_{0}_{(1)}$)) \Rightarrow

• (
$$c^{(0)}$$
 , $c^{(1)}$) = (Enc(K, $m_1^{(0)}$), Enc(K, $m_1^{(1)}$))
 $\Rightarrow c^{(0)} \neq c^{(1)}$ (by correctness)

Generalize to Randomized Encryption

$$(c^{(0)}, c^{(1)}) = (Enc(K, m), Enc(K, m))$$

$$Pr[c^{(0)} = c^{(1)}]$$
?

- Fix **k**
- Conditioned on k, $c^{(0)}$, $c^{(1)}$ are two independent samples from same distribution Enc(k, m)

Lemma: Given any distribution D over a finite set X, $Pr[Y=Y': Y\leftarrow D, Y'\leftarrow D] \ge 1/|X|$

• Therefore, $Pr[c^{(0)} = c^{(1)}]$ is non-zero

Generalize to Randomized Encryption

Let (Enc, Dec) be stateless, deterministic

Let
$$\mathbf{m}_0^{(0)} = \mathbf{m}_0^{(1)}$$

Let $\mathbf{m}_1^{(0)} \neq \mathbf{m}_1^{(1)}$

Consider distributions of encryptions:

• (
$$c^{(0)}$$
 , $c^{(1)}$) = ($Enc(K, m_0^{(0)})$, $Enc(K, m_0^{(1)})$)
$$\Rightarrow Pr[c^{(0)} = c^{(1)}] > 0$$
• ($c^{(0)}$, $c^{(1)}$) = ($Enc(K, m_1^{(0)})$, $Enc(K, m_1^{(1)})$)
$$\Rightarrow Pr[c^{(0)} = c^{(1)}] = 0$$

What do we do now?

Tolerate tiny probability of distinguishing

• If $Pr[c^{(0)} = c^{(1)}] = 2^{-128}$, in reality never going to happen

How small is ok?

• Practitioner: **2**⁻⁸⁰, **2**⁻¹²⁸, or maybe **2**⁻²⁵⁸

• Theorist: ????

Big O Notation Recap

O(
$$f(\lambda)$$
) = { $g(\lambda)$: $\exists c, \lambda_0, \forall \lambda > \lambda_0 g(\lambda) \le cf(\lambda)$ }
 $\Omega(f(\lambda)) = {g(\lambda)$: $\exists c, \lambda_0, \forall \lambda > \lambda_0 g(\lambda) \ge cf(\lambda)$ }
 $\Theta(f(\lambda)) = O(f(\lambda)) \cap \Omega(f(\lambda))$
o($f(\lambda)) = O(f(\lambda)) \setminus \Omega(f(\lambda))$
 $\omega(f(\lambda)) = \Omega(f(\lambda)) \setminus O(f(\lambda))$

Notation abuse:

$$g(\lambda) = O(f(\lambda))$$
 means $g(\lambda) \in O(f(\lambda))$

Polynomial

Def: $f(\lambda)$ is polynomially bounded if $f(\lambda) \subseteq O(\lambda^c)$ for some constant c

- Sometimes will just say "f(λ) is polynomial"
- Equivalent def: $log(f(\lambda)) = O(log \lambda)$
- Set of polynomial functions: **n**O(1)

Def: $f(\lambda)$ is inverse polynomial if $f(\lambda) \subseteq \Omega(\lambda^{-c})$ for some constant c

Def: $f(\lambda)$ is super polynomial if $f(\lambda)$ $\notin O(\lambda^c)$ for any constant c

Negligible

Def: $f(\lambda)$ is **negligible** if $f(\lambda) \subseteq O(\lambda^{-c})$ for any constant c

- Equivalent def: $log(1/f(\lambda)) \in \omega(log \lambda)$
- Set of negligible functions: 2-ω(log λ)

Negligible Function Examples

Negligible:

- $c^{-\lambda}$ for any constant c
- $\lambda^d c^{-\lambda}$ for any constants c,d
- λ^{-log λ} (for any logarithm base)

Not negligible:

- λ⁻²
- 1/log λ
- $\lambda^{-2}/\log \lambda$

Negligible?

Is the following function negligible?

$$f(\lambda) = 2^{-\lambda^{1/2}}$$

Yes: $log(1/f(\lambda)) = \lambda^{1/2} \in \omega(log \lambda)$

Negligible?

Is the following function negligible?

$$f(\lambda) = 2^{-\lambda}$$
 if λ odd, $1/\lambda$ if λ even

No: $f(\lambda) \notin O(\lambda^{-2})$

Security Parameter

System parameterized by security parameter λ

- Represents security level of system
- System parameters allowed to grow poly in λ
 - i.e. |ctxt|, $|key| = O(\lambda^c)$ for some constant c
- Adversary distinguishing advantage **negligible** in λ
- Idea: poly = tractable, negl/superpoly = intractable

Encryption with Security Parameter

Syntax:

- Key space K_λ
- Message space M (sometimes depends on λ)
- Ciphertext space C_{λ}
- Enc: $K_{\lambda} \times M \rightarrow C_{\lambda}$
- Dec: $K_{\lambda} \times C_{\lambda} \rightarrow M$

Statistical Secrecy

Definition: A scheme (Enc,Dec) has statistical secrecy for n messages if \exists negligible function ε s.t. \forall two sequences of messages $(m_0^{(i)})_{i\in[n]}$, $(m_1^{(i)})_{i\in[n]} \in M^n$ $\Delta \big[\big(\text{Enc}(K_\lambda, \, m_0^{(i)} \,) \big)_{i\in[n]}, \big(\text{Enc}(K_\lambda, \, m_1^{(i)} \,) \big)_{i\in[n]} \big] < \varepsilon(\lambda)$

Ex:

$$M_{\lambda} = C_{\lambda} = \mathbb{Z}_{p}$$
 (p a prime of size 2^{λ})
 $K_{\lambda} = \mathbb{Z}_{p}^{*} \times \mathbb{Z}_{p}$
 $Enc((a,b), m) = (am + b) \mod p$
 $Dec((a,b), c) = (c-b)/a \mod p$

Q: Is this statistically secure for two messages?

Ex:

$$M_{\lambda} = C_{\lambda} = \mathbb{Z}_{p}$$
 (p a prime of size 2^{λ})
 $K_{\lambda} = \mathbb{Z}_{p}^{*} \times \mathbb{Z}_{p}$
 $Enc((a,b), m) = (am + b) \mod p$
 $Dec((a,b), c) = (c-b)/a \mod p$

Attack:

- $m_0^{(0)} = m_0^{(1)} = 0$, $m_1^{(0)} = 0$, $m_1^{(1)} = 1$
- $A(c^{(0)},c^{(1)}) = 1$ iff $c^{(0)}=c^{(1)}$
- Advantage: 1 (non-negligible)

Example

Ex:

$$M_{\lambda} = \mathbb{Z}_p$$
 (**p** a prime of size 2^{λ})
 $C_{\lambda} = \mathbb{Z}_p^2$
 $K_{\lambda} = \mathbb{Z}_p^2$
Enc((a,b), m) = (r, (ar+b) + m)
 $Dec((a,b), (r,c)) = c - (ar+b)$

Q: Is this statistically secure for two messages?

(d+1)-time Secure Encryption

Ex:

$$M_{\lambda} = \mathbb{Z}_{p}$$
 (p a prime of size 2^{λ})
 $C_{\lambda} = \mathbb{Z}_{p}^{2}$
 $K_{\lambda} = \{\text{degree d polynomials over } \mathbb{Z}_{p}\}$
 $\text{Enc}(P, m) = (r, P(r) + m)$
 $\text{Dec}(P, (r,c)) = c - P(r)$

Theorem: This scheme is secure for up to **d+1** messages

What is the distribution (Enc(K_{λ} , $m^{(i)}$)) $_{i \in [d+1]}$?
• First, fix ($r^{(i)}$) $_{i \in [d+1]}$

```
Claim: If the \mathbf{r}^{(i)} are distinct (\mathbf{r}^{(i)} \neq \mathbf{r}^{(j)}) for any \mathbf{i} \neq \mathbf{j}, then  (\mathbf{P}(\mathbf{r}^{(i)}))_{i \in [d+1]} \stackrel{d}{=} \mathbb{Z}_p^{d+1}
```

Corollary: If the $\mathbf{r}^{(i)}$ are distinct $(\mathbf{r}^{(i)} \neq \mathbf{r}^{(j)})$ for any $\mathbf{i} \neq \mathbf{j}$, then $(\mathbf{P}(\mathbf{r}^{(i)}) + \mathbf{m}^{(i)})_{i \in [d+1]} \stackrel{d}{=} \mathbb{Z}_p^{d+1}$

Proof of Claim:

- Fix distinct (r⁽ⁱ⁾)_{i∈[d+1]}
- Fix tuple (y⁽ⁱ⁾)_{i∈[d+1]}
- Exactly one P such that P(r(i)) = y(i) for all i
- Total number of polynomials: p^{d+1}
- Pr[(P(r(i))) $_{i \in [d+1]} = (y^{(i)})_{i \in [d+1]}$]= 1/p^{d+1}
- Therefore ($P(r^{(i)})$)_{$i \in [d+1]$} $\stackrel{d}{=} \mathbb{Z}_p^{d+1}$

What is the distribution (Enc(K_{λ} , $m^{(i)}$)) $_{i \in [d+1]}$?
• First, fix ($r^{(i)}$) $_{i \in [d+1]}$

```
Claim: If the \mathbf{r}^{(i)} are distinct (\mathbf{r}^{(i)} \neq \mathbf{r}^{(j)}) for any \mathbf{i} \neq \mathbf{j}, then  (\mathbf{P}(\mathbf{r}^{(i)}))_{i \in [d+1]} \stackrel{d}{=} \mathbb{Z}_p^{d+1}
```

Corollary: If the $\mathbf{r}^{(i)}$ are distinct $(\mathbf{r}^{(i)} \neq \mathbf{r}^{(j)})$ for any $\mathbf{i} \neq \mathbf{j}$, then $(\mathbf{P}(\mathbf{r}^{(i)}) + \mathbf{m}^{(i)})_{i \in [d+1]} \stackrel{d}{=} \mathbb{Z}_p^{d+1}$

Lemma: $\Delta(D_1,D_2) \leq \Pr[bad|D_1] + \Pr[bad|D_2] + \Delta(D_{1,good},D_{2,good})$

Where:

- "bad" is some event
- Pr[bad|D_b] is probability "bad" when sampling from D_b
- D_{b,good} is the distribution D_b conditioned on not "bad"

Proof of Lemma

$$\begin{split} \Delta(D_{1},D_{2}) &= \Sigma_{x} \middle| \ \Pr[D_{1}=x] - \Pr[D_{2}=x] \middle| \\ &= \Sigma_{x:bad} \middle| \ \Pr[D_{1}=x] - \Pr[D_{2}=x] \middle| \\ &+ \Sigma_{x:good} \middle| \ \Pr[D_{1}=x] - \Pr[D_{2}=x] \middle| \\ &\leq \Sigma_{x:bad} \middle| \ \Pr[D_{1}=x] \middle| + \Sigma_{x:bad} \middle| \ \Pr[D_{2}=x] \middle| \\ &+ \Sigma_{x:good} \middle| \ \Pr[D_{1}=x] - \Pr[D_{2}=x] \middle| \\ &\leq \Pr[bad|D_{1}] + \Pr[bad|D_{2}] + \Delta(D_{1,good},D_{2,good}) \end{split}$$

Back to Security Proof

Goal: bound
$$\Delta$$
 ($P(r^{(i)})+m^{(i)}$) $_{i\in[d+1]}$, \mathbb{Z}_p^{d+1})

Define "**bad**" to be that the $r^{(i)}$ are not distinct

- Conditioned on "good", $\Delta=0$
- So using previous lemma Δ ≤ 2Pr[bad]

Lemma: $Pr[bad] \leq (d+1)^2/2p$

Lemma:
$$Pr[bad] \le (d+1)^2/p$$
 $Pr[bad] = Pr[r^{(1)}=r^{(2)} \text{ or } r^{(1)}=r^{(3)} \text{ or } ... \text{ or } r^{(1)}=r^{(d+1)}$

or $r^{(2)}=r^{(3)} \text{ or } ...$]

 $\le Pr[r^{(1)}=r^{(2)}] + Pr[r^{(1)}=r^{(3)}] + ... + Pr[r^{(1)}=r^{(d+1)}]$
 $+ Pr[r^{(2)}=r^{(3)}] + ... \text{ (Union Bound)}$
 $= (1/p) {d+1 \choose 2}$
 $\le (d+1)^2/2p$

Back to Security Proof

Goal: bound
$$\Delta$$
 ($P(r^{(i)})+m^{(i)}$) $_{i\in[d+1]}$, \mathbb{Z}_p^{d+1})

Define "**bad**" to be that the $r^{(i)}$ are not distinct

- Conditioned on "good", $\Delta=0$
- So using previous lemma △ ≤ Pr[bad]

Lemma:
$$Pr[bad] \leq (d+1)^2/2p$$

• So $\Delta \leq (d+1)^2/p$

Finishing up the proof

$$\begin{split} \Delta \big[& \left(\mathsf{Enc}(\mathsf{K}_{\lambda}, \, \mathsf{m}_0^{(i)} \, \right) \big)_{i \in [n]} \, , \left(\mathsf{Enc}(\mathsf{K}_{\lambda}, \, \mathsf{m}_1^{(i)} \, \right) \big)_{i \in [n]} \, \big] \\ & \leq \Delta \big[\left(\mathsf{Enc}(\mathsf{K}_{\lambda}, \, \mathsf{m}_0^{(i)} \, \right) \big)_{i \in [n]} \, , \, \mathbb{Z}_p^{d+1} \, \big] \\ & + \Delta \big[\left(\mathsf{Enc}(\mathsf{K}_{\lambda}, \, \mathsf{m}_1^{(i)} \, \right) \big)_{i \in [n]} \, , \, \mathbb{Z}_p^{d+1} \, \big] \\ & \leq 2(d+1)^2/p \leq 2(d+1)^2/2^{\lambda} \end{split}$$

Summary

Stateful encryption is hard to manage Stateless encryption cannot be perfectly secure for multiple messages

Therefore, use statistical security

Unfortunately, for our example, total number of messages bounded by key length

Really want unbounded number of messages

Next Time

Bound on message length/number of messages necessary for our security definitions

Computational security: security against computationally bounded adversaries

- Allows for keys that are very small (e.g. 128 bits)
- Can encrypt arbitrary number of messages of arbitrary length
- However, cannot prove security unconditionally