Notes from CS 6260 (Applied Cryptography)
Georgia Tech, Fall 2012

Christopher Martin
chris.martin@gatech.edu

1 Symmetric cryptography scheme

Key space	$\mathcal{K}$
Message space	$\mathcal{M}$
Cypher space	$\mathcal{C}$
Key generator	$\mathsf{Gen} : \phi \to \mathcal{K}$
Encryption function	$\mathsf{Enc} : \{\mathcal{K} \times \mathcal{M}\} \to \mathcal{C}$
Decryption function	$\mathsf{Dec} : \{\mathcal{K} \times \mathcal{C}\} \to \mathcal{M}$

2 Information theoretic security

Information theoretic security repels even resource-unbounded attackers. Shannon secrecy and perfect secrecy are equivalent definitions of information theoretic security for symmetric cryptography schemes.

Shannon secrecy A scheme is Shannon-secret with respect to the distribution D over $\mathcal{M}$ iff the ciphertext reveals no additional information about the message.

$$\forall\, M \in \mathcal{M},\, C \in \mathcal{C} : \Pr_{\substack{k \leftarrow \mathsf{Gen} \\ m \in D}} \big[\, m = M \,|\, \mathsf{Enc}_k(m) = C \,\big] = \Pr_{m \in D} \big[\, m = M \,\big]$$

Perfect secrecy A scheme is perfectly secret iff the distributions of ciphertexts for any two messages are identical.

$$\forall\, M_1, M_2 \in \mathcal{M},\, C \in \mathcal{C} : \Pr_{K_1 \leftarrow \mathsf{Gen}} \big[\, \mathsf{Enc}_{K_1}(M_1) = C \,\big] = \Pr_{K_2 \leftarrow \mathsf{Gen}} \big[\, \mathsf{Enc}_{K_2}(M_2) = C \,\big]$$

This model considers only a single message and ciphertext, so although a one-time pad is perfectly secret, a “two-time pad” is not.

Theorem 1. Perfect secrecy $\Rightarrow |\mathcal{K}| \geq |\mathcal{M}|$.

Proof. If not, $\exists$ 2 messages with different probabilities of encrypting to the same cypertext. $\square$

3 Pseudo-random functions

Uniformly random function U is a random variable chosen uniformly from the set of all functions $\{0,1\}^m \to \{0,1\}^n$.

Pseudo-random function A PRF belongs to a family of functions $F : \{0,1\}^\ell \times \{0,1\}^m \to \{0,1\}^n$. Write $F_k(\cdot)$ to denote $F(k, \cdot)$.

Distinguishing advantage Consider an adversary $\mathcal{A}$ who knows F, having oracle access to F_k where k was chosen uniformly at random, trying to distinguish the oracle's responses from a random function. The distinguishing advantange of $\mathcal{A}$ against F is

$$\mathrm{Adv}_F^{\mathrm{prf}}(\mathcal{A}) \equiv \Pr_{k \in \{0,1\}^\ell} \left[\mathcal{A}^{F_k} \text{ accepts} \right] - \Pr_U \left[\mathcal{A}^U \text{ accepts} \right] .$$

In time $O(t)$, we can brute-force t keys to get advantage $t/2^\ell$.

(t,q)-bounded adversary

t	Running time
q	Number of queries

(t,q,ε)-secure PRF F is (t,q,ε)-secure iff $\forall$ (t,q)-bounded $\mathcal{A}$,

$$\mathrm{Adv}_F^{\mathrm{prf}}(\mathcal{A}) \leq \varepsilon .$$

Examples of reasonable constants

t	2^{128}
q	2^{64} or 2^{32}
ε	2^{-128}

Existence The existence of secure PRFs has not been proven, but there are some functions that have never been broken and are widely assumed to be PRFs.

4 Reduction

Karp (many-to-one) reduction Reduction from A to B transforms an instance of A to an instance of B.

Cook (Turing) reduction Reduction from A to B solves A using a subroutine that solves B.

Key recovery security F is (t, q, ε)-kr-secure iff $\forall$ (t, q)-bounded $\mathcal{A}$,

$$\mathrm{Adv}_F^{\mathrm{kr}} \equiv \Pr_{k \in \{0,1\}^\ell} \left[\mathcal{A}^{F_k(\cdot)} \text{ outputs } k \right] \leq \varepsilon \ .$$

Theorem 2. If F is a (t, q, ε)-secure PRF for $q < 2^m$, then F is (t', q', ε')-kr-secure for $t' \approx t$, $q' = q - 1$, $\varepsilon' = \epsilon + 2^{-n}$.

Proof. Cook reduction. For any kr-adversary $\mathcal{A}'$ running in time t' and making $q' < 2^m$ queries, let $\mathcal{A}$ be the PRF adversary:

> $k' \leftarrow \mathcal{A}'(\mathcal{O})$
> $x \leftarrow$ a value that $\mathcal{A}'$ did not query with
> $y \leftarrow \mathcal{O}(x)$
> Accept iff $y = F_{k'}(x)$

$\mathcal{A}$ runs in time $t \approx t'$ and makes $q = q' + 1$ queries.

$$\mathrm{Adv}_F^{\mathrm{prf}}(\mathcal{A}) \geq \mathrm{Adv}_F^{\mathrm{kr}}(\mathcal{A}') - 2^{-n} \ . \qquad \square$$

Example PRF construction For PRF $F : \{0,1\}^\ell \times \{0,1\}^n \to \{0,1\}^n$, $F'_k(x) \equiv F_k(F_k(x)) \| F_k(\overline{F_k(x)})$

Theorem 3. F' is $(t, \frac{q}{3}, \varepsilon + \frac{q^2}{2^n})$-secure.

Proof. Let $\mathcal{A}'$ be an attacker on F'. Define $\mathcal{A}$ as:

> $\mathcal{O}' \equiv \mathcal{O}(\mathcal{O}(x)) \| \mathcal{O}(\overline{\mathcal{O}(x)})$ (done with 3 queries to $\mathcal{O}$)
> Accept iff ${\mathcal{A}'}^{\mathcal{O}'}$ accepts

$\mathcal{O}'(x)$ simulates F' perfectly, so $\Pr_k \left[\mathcal{A}^{F_k} \right] = \Pr_k \left[{\mathcal{A}'}^{F'_k} \right]$.

$\mathcal{O}'$ does not simulate U perfectly, but it is close. We have independence as long as all of the $\mathcal{O}(x)$, $\overline{\mathcal{O}(x)}$ are distinct. Using union bound, this probability $\leq \frac{q^2}{2^n}$ $\square$

5 Pseudo-random permutations

In a permutation family $F : \{0,1\}^\ell \times \{0,1\}^n \to \{0,1\}^n$, every F_k is bijective.

A secure PRP is computationally indistinguishable from a uniformly random permutation.

Some well-known PRPs

DES	$\ell = 56$	$n = 64$
AES_{128}	$\ell = 128$	$n = 128$
AES_{192}	$\ell = 192$	$n = 128$

Strong PRP / block cipher Attackers with oracle access to both F and F^{-1} have small advantage.

$$\mathrm{Adv}_F^{\mathrm{sprp}} \equiv \Pr_k \left[\mathcal{A}^{F_k, F_k^{-1}} \text{ accepts} \right] - \Pr_P \left[\mathcal{A}^{P, P^{-1}} \text{ accepts} \right] \leq \varepsilon$$

PRF/PRP switching lemma If G is a (t, q, ε)-secure PRP (not necessarily strong), then F is a $(t, q, \varepsilon + \frac{q^2}{2^{n+1}})$-secure PRF.

6 Secure symmetric encryption

Perfect secrecy is impossible where $m > \ell$, but computational security is possible with pseudorandom objects.

Electronic code block (ECB) Suppose F is a secure PRP $\{\mathtt{0},\mathtt{1}\}^\ell \times \{\mathtt{0},\mathtt{1}\}^n \to \{\mathtt{0},\mathtt{1}\}^n$ with F and F^{-1} efficiently computable.

Gen	$k \leftarrow \{\mathtt{0},\mathtt{1}\}^\ell$
Enc	$M' \leftarrow$ Pad message M with 1 and some 0s to a multiple of n. Break M' into n-bit blocks $m_0, m_1, \ldots$ Apply F_k to each of the $\{m\}$
Dec	Apply F_k' to each of the $\{m\}$

Repeated blocks give repeated ciphertext. Never use ECB.

Security model Adversary, seeing all cipthertexts and having oracle access to Enc_k, learns nothing about plaintexts (except message length, which is unavoidable).

$SE = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is (t, σ, ε)-IND-CPA secure ("indistinguishable under chosen-plaintext attack") iff $\forall$ (t, σ)-bounded $\mathcal{A}$,

$$\mathrm{Adv}^{\mathrm{indcpa}}_{SE}(\mathcal{A}) \equiv \Pr_{k \leftarrow \mathsf{Gen}}\left[\,\mathcal{A}^{L_k} \text{ accepts}\,\right] - \Pr_{k \leftarrow \mathsf{Gen}}\left[\,\mathcal{A}^{R_k} \text{ accepts}\,\right] ,$$

$$L_k(m, m') \equiv \mathsf{Enc}_k(m) \text{ if } |m| = |m'| \text{ else } \perp ,$$

$$R_k(m, m') \equiv \mathsf{Enc}_k(m') \text{ if } |m| = |m'| \text{ else } \perp ,$$

t is the running time, and σ total length of all message queries.

Equivalent definition: Enc_k is computationally indistinguishable from a zero-encrypting oracle $Z_k \equiv \mathsf{Enc}_k(\mathsf{0}^m)$.

Query repetition Enc in an IND-CPA-secure scheme should not always return the same ciphertext for multiple encryptions of the same message. This attack has advantage 1 against any deterministic and stateless scheme:

$c \leftarrow \mathcal{O}(\langle 0 \rangle, \langle 0 \rangle)$
$c' \leftarrow \mathcal{O}(\langle 0 \rangle, \langle 1 \rangle)$
Accept iff $c = c'$

7 Block cipher modes

Stateful counter mode (CTRS) Let F be a PRF with $m = n$.

Gen | $k \leftarrow \{\mathsf{0}, \mathsf{1}\}^{\ell}$, $\mathit{counter} \leftarrow 0$

Enc | echo $\mathit{counter}$
for each message block m:
  echo $F_k(\mathit{counter}) \oplus m_i$
  increment $\mathit{counter}$

CTRS is not used much, because preserving *counter* is difficult.

Theorem 4. If F is a (t, q, ε)-secure PRF, then CTRS(F) is $(t' \approx t, qn, 2\varepsilon)$-IND-CPA secure.

Proof. We will show using a hybrid argument that $\forall$ (t', σ)-bounded $\mathcal{A}'$ against CTRS(F) where $\sigma \le n\,2^m$, there is a $(t \approx t', q = \sigma/n)$-bounded attacker $\mathcal{A}$ attacking F such that $\mathrm{Adv}^{\mathrm{indcpa}}_{\mathrm{CTRS}(F)}(\mathcal{A}') \le 2\,\mathrm{Adv}^{\mathrm{prf}}_{F}(\mathcal{A})$.

Given $\mathcal{O}$ that is either F or U:

$$\mathcal{A} \equiv \mathcal{A}_L \equiv \left| \begin{array}{l} \mathit{counter} \leftarrow 0 \\ \mathcal{O}'(m, m') \equiv \left| \begin{array}{l} \text{If } |m| = |m'|\text{, return } \bot \\ \text{Split } m \text{ into blocks } m_0, m_1, \ldots, m_{t-1} \\ y_i \leftarrow \mathcal{O}(\mathit{counter} + i) \ \forall\ i \in [0, t) \\ \text{Return } \mathit{counter} \| \text{join}_i(m_i \oplus y_i) \\ \mathit{counter} \leftarrow \mathit{counter} + t \end{array} \right. \\ \text{Accept iff } \mathcal{A}'^{\mathcal{O}'} \text{ accepts} \end{array} \right.$$

Also define $\mathcal{A}_R$ similarly using m' instead of m.

$\mathcal{A}_L^{F_k}$ perfectly simulates L_k to $\mathcal{A}'$.

$\mathcal{A}_L^U$ does not simulate R_k, but it does simulate an oracle \$:

$$\$(m, m') \equiv \left| \begin{array}{l} \text{If } |m| = |m'|\text{, return } \bot \\ \text{Return } \mathit{counter} \| [\text{random bits}] \\ \mathit{counter} \leftarrow \mathit{counter} + \text{number of blocks} \end{array} \right.$$

$$\begin{aligned} P_\ell &= \Pr_k[\mathcal{A}_L^{F_k}] = \Pr_k[\mathcal{A}'^{L_k}] \\ P_r &= \Pr_k[\mathcal{A}_R^{F_k}] = \Pr_k[\mathcal{A}'^{R_k}] \\ P_\$ &= \Pr_k[\mathcal{A}_L^U] = \Pr_k[\mathcal{A}_R^U] = \Pr_k[\mathcal{A}'^{\$}] \end{aligned}$$

$$\begin{aligned} \text{Adv}^{\text{indcpa}}_{\text{CTRS}(F)}(\mathcal{A}') &= |P_\ell - P_r| \\ &\leq |(P_\ell - P_\$) + (P_r - P_\$)| \qquad \text{(triangle inequality)} \\ &\leq \varepsilon + \varepsilon = 2\varepsilon \end{aligned}$$

□

Counter modes

CTRS	One global counter	$\text{adv}^{\text{indcpa}} \leq 2\,\text{adv}^{\text{prf}}$
CTR\$	Random IV for each message	$\text{adv}^{\text{indcpa}} \leq 2\,\text{adv}^{\text{prf}} + q^2/2^n$
CTR\$\$	Random IV for each block	

Cipher block chaining (CBC) $C_0 = \text{IV}$, $C_i = F_k(C_{i-1} \oplus m_i)$

Dec requires being able to calculate F^{-1}.

If F is a (t, q, ε)-secure PRF, then CBC$[F]$ is $(\approx t, \sigma = qn, 2\varepsilon + q^2/2^n)$-ind-cpa-secure. The proof requires showing that for U, all inputs to U are distinct (minus a birthday term).

8 Message authentication code (MAC)

Alice sends message m and $t \leftarrow \mathsf{Tag}_k(m)$. Eve intercepts (m, t) and delivers (m', t') to Bob. Bob runs $\mathsf{Ver}_k(m', t')$.

$$\mathsf{Ver}_k \text{ returns } \begin{cases} m & \text{if } t' \text{ is a valid tag } (\mathsf{Ver}_k \text{ “accepts”}) \\ \perp & \text{otherwise } (\mathsf{Ver}_k \text{ “rejects”}) \end{cases}$$

Eve has access to a Tag_k oracle and can make many attempts on Ver. Eve “wins” if Ver accepts on an m' not previously queried to Tag_k.

Conerns ignored by this model

Dropped messages

Replay attacks (“freshness” of messages)

Message sequence

Unforgeability under chosen message attack

$$\mathrm{Adv}^{\mathrm{ufmca}}_{\mathrm{MAC}}(\mathcal{A}) \equiv \Pr_{k \leftarrow \mathsf{Gen}} \left[\mathcal{A}^{\mathsf{Tag}_k, \mathsf{Ver}_k} \text{ “wins” } \right]$$

MAC is $(t, q_t, q_v, \varepsilon)$-uf-cma-secure iff advantage of an attacker bounded by time t, number of Tag queries q_t, and number of Ver queries q_v is less than ε.

Examples of reasonable constants

t	2^{80} or 2^{128}
q_v, q_t	2^{40} or 2^{56}
ε	2^{-40} or 2^{-56}

Brute-force MAC attacks

Key search: Get a few oracle tags, and guess k. $\mathrm{Adv} = t/2^{\ell}$.

Tag search: $\mathrm{Adv} = t/2^s$ where s is the tag length.

PRF-based MAC $\mathsf{Tag}_k \equiv F_k$

$\forall$ (t, q_t, q_v)-bounded $\mathcal{B}$, $\exists$ $(\approx t, q_t + q_v)$-bounded $\mathcal{A}$ such that

$$\mathrm{Adv}^{\mathrm{ufcma}}_{\mathrm{PRFMAC}[F]}(\mathcal{B}) \leq \mathrm{Adv}^{\mathrm{prf}}_{F}(\mathcal{A}) + q_v/2^n .$$

CBC-MAC For a fixed t, and $F : \{0,1\}^{nt} \to \{0,1\}^n$, CBC-MAC$[F]$ is secure, losing $(qt)^2/2^n$ advantage from that of F.

Cipher-based MAC (CMAC) Adds an extra step to the end of CBC-MAC to make it secure for arbitrary-length messages.

Precompute k_1, $k_2 \in \{0, 1\}^n$ using $F_k(0^m)$.

$$m'_t \leftarrow \begin{cases} m'_t \oplus k_1 & : |m'_t| = n \\ m' \| 000 \ldots \oplus k_2 & : |m'_t| < n \end{cases}$$

Run $m_1 \| \ldots \| m_t$ through CBC-MAC.

9 Combining authenticity and privacy

Integrity of ciphertexts (INT-CTXT) $\mathsf{Dec}_k(c)$: returns decryption of c, or $\perp$ if c is invalid.

$SE = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is INT-CTXT secure iff $\forall$ bounded $\mathcal{A}$,

$$\mathrm{Adv}_{SE}^{\text{int-ctxt}}(\mathcal{A}) \equiv \Pr_{k \leftarrow \mathsf{Gen}} \left[\mathcal{A}^{\mathsf{Enc}_k, \mathsf{Dec}_k} \text{ wins} \right] < \varepsilon \ .$$

UF-CMA-security does not necessarily give INT-CTXT security. For example: If the output of Tag has a spurious bit that is ignored by Ver. So we require a stronger condition:

Strong unforgeability (SUF-CMA) Winning is redefined as: Ver_k accepts (m', t') that was not previously a query/answer pair to Tag_k.

Bad idea: Encrypt-and-tag $\mathsf{AEnc} \equiv \mathsf{EEnc}_{k_e}(m) \| \mathsf{Tag}_{k_m}(m)$. The tag could reveal information about m.

Bad idea: Tag-then-encrypt $\mathsf{AEnc} \equiv \mathsf{EEnc}_{k_e}(m \| \mathsf{Tag}_{k_m}(m))$. The ciphertext might be forgeable (for example, if EEnc appends a spurious bit).

Good idea: Encrypt-then-tag $\mathsf{AEnc} \equiv \mathsf{EEnc}_{k_e}(m) \| \mathsf{Tag}_{k_m}(\mathsf{EEnc}_{k_e}(m))$.

Indistinguishability under chosen ciphertext attack (IND-CCA) IND-CPA $\wedge$ INT-CTXT $\Rightarrow$ IND-CCA.