Lecture 13: Left or Right Security

The focus of this lecture is Left-or-Right (LOR) security, which was defined at the end of the last lecture. We will discuss some of the implications of LOR security in terms of other concepts of security, as well as some systems that are provably LOR secure. First we review the definition of LOR security:

Definition 1 (Left or Right Security) *Let $\S_i(x_0, x_1) = x_i$. Then an ecryption system E is (t, q, μ, ϵ)-secure in the LOR sense against CPA if for all adversaries A running in time t and making q queries of total length at most μ,*

$$\left|\mathbf{Pr}_{i,k}[A^{E_k \circ \S_0} = 1] - \mathbf{Pr}_{i,k}[A^{E_k \circ \S_1} = 1]\right| \leq \epsilon.$$

Note that we require that $|x| = |y|$ for oracle call (x, y).

Informally, this definition means that A cannot tell, except with low probability, whether on input (x, y), the oracle returns the ecryption of x or the encryption of y. LOR security is a very strong notion of security. We will show that it implies security against key recovery attacks and security against CPA.

Theorem 1 *LOR security implies security against key recovery attacks.*

Proof: If A^{E_k} is a key-recovery attack, we create $B^{E_k \circ \S_i}$ to be an LOR attack. Let $\mathcal{O} = E_k \circ \S_i$. Define $B^{\mathcal{O}}$ as follows:

1. Let $k' \leftarrow A^f$ where $f(m) = \mathcal{O}(m, m)$, so k' is the decryption key.

2.Let $c \leftarrow \mathcal{O}(R, R')$ where $R, R' \in_R \mathcal{M}$, i.e. R and R' are messages drawn at random from the same message space $\mathcal{M}$.

3. If $D_{k'}(c) = R$, output 1; else output 0.

When A succeeds in recovering the correct key, then B almost always succeds in distinguishing whether the left or right message is being encrypted by $\mathcal{O}$. The only case when B may fail though A is successful is when $R = R'$. A big enough message space eliminates this problem.

□

Note that more generally, A just needs to decrypt for the attack to be successful; it doesn't actually need to recover the key. Almost the same proof works to show that LOR security implies normal CPA security.

Theorem 2 *If (E, D) is (t, q, μ, ϵ)-secure in the LOR sense against CPA, then $\mathbf{Pr}_{M \leftarrow \mathcal{M}}[A^{E_k}(E_k(M)) = M] \leq \epsilon + (1/|\mathcal{M}|)$ for any adversary A in time $t - o(1)$ with $q - 1$ queries.*

Proof: Define $B^{\mathcal{O}}$ to be an LOR-attacker as follows:

1. Pick R, R' uniformly at random from $\mathcal{M}$. Let $c = \mathcal{O}(R, R')$, so c is the encryption of either R or R'.

2. Let $M \leftarrow A^f(c)$ where $f(m) = \mathcal{O}(m, m)$.
3. If $M = R$ output 1; else output 0.

Suppose A is a good attacker, i.e. suppose $\mathbf{Pr}_{M \leftarrow \mathcal{M}}[A^f(E_k(M)) = M] > \epsilon + 1/|\mathcal{M}|$. Then $\mathbf{Pr}[B^{E_k \circ \S_0} = 1] > \epsilon + 1/|\mathcal{M}|$, and $\mathbf{Pr}[B^{E_k \circ \S_1} = 1] = \mathbf{Pr}[R = R'] = 1/|\mathcal{M}|$. From this we conclude that (E, D) is not ϵ-secure in the LOR sense after all since

$$\left| \mathbf{Pr}[B^{E_k \circ \S_0} = 1] - \mathbf{Pr}[B^{E_k \circ \S_1} = 1] \right| > \epsilon.$$

□

Now we will show that counter (CTR) mode is LOR-secure. First we begin with simplified counter mode (SIMPLE-CTR).

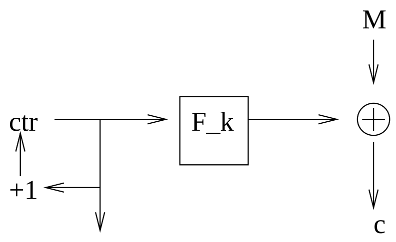

Figure 1: SIMPLE-CTR

Theorem 3 *If F is a (t, q, ϵ)-secure PRF, $F_k : \{0,1\}^l \to \{0,1\}^m$, then SIMPLE[$F_k$] is $(t, q, mq, 2\epsilon)$-secure against LOR-CPA, if $q \le 2^l$.*

(Note that we need $q \le 2^l$ to be sure that the counter doesn't wrap.) We give an information theoretic proof by first showing that SIMPLE[F_k] is secure if we use a truly random function F_k and then show that it remains secure if we replace this function with a PRF.

Lemma 4 *If $R : \{0,1\}^l \to \{0,1\}^m$ is a random function, then SIMPLE[R] is $(\infty, q, mq, 0)$-secure against LOR-CPA, if $q \le 2^l$.*

Proof: This should be reminiscent of the one-time pad, which is perfectly secure.

Consider an adversary A for LOR-CPA. A sends queries (M_0, M_1) to the oracle $R \circ \S_i$ and tries to guess if the output of the oracle is the encryption of M_0 or M_1 and behave differently accordingly. What kind of information does A get about i from interacting with the oracle? The counter does not depend on i, and R does not depend on i. Therefore the XOR mask $R(\text{ctr})$ is independent of i, and the distribution of $R(\text{ctr}) \oplus M_i$ is the same whether $i = 1$ or $i = 0$. So

$$\mathbf{Pr}[A^{R \circ \S_0} = 1] = \mathbf{Pr}[A^{R \circ \S_1} = 1].$$

□

Proof: (of Theorem 3) Now we use a PRF instead of R. Let $p = \mathbf{Pr}_R[A^{\text{SIMPLE}[R] \circ \S_0} = 1] = \mathbf{Pr}_R[A^{\text{SIMPLE}[R] \circ \S_1} = 1]$, and let $q_i = \mathbf{Pr}_k[A^{\text{SIMPLE}[F_k] \circ \S_i} = 1]$. If F is a secure

PRF, then $|q_i - p| \leq \epsilon$, since otherwise we could use, for example, $A^{\mathrm{SIMPLE}[F_k]\circ\S_0}$ to distinguish F_k from R. So then $|q_1 - q_0| \leq 2\epsilon$. $\square$

Now we consider stateful counter mode (S-CTR). We allow message lengths to vary, although they are bounded. Note that in an example where the first message has five blocks, the counter would then start at five for the second message.

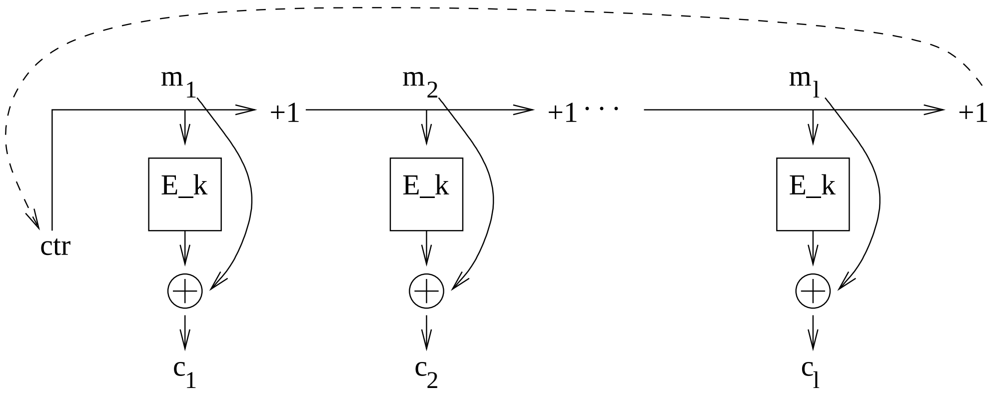

Figure 2: S-CTR

Theorem 5 *If $F_k : \{0,1\}^l \to \{0,1\}^L$ is a $(t, \mu/L, \epsilon)$-secure PRF, then S-CTR[F_k] is $(t, q, \mu, 2\epsilon)$-secure against LOR-CPA, if $\mu/L \leq 2^l$.*

Once again, we need $\mu/L \leq w^l$ so that the counter does not wrap. If we imagine replacing F_k with a truly random function R, then we would need all inputs to R to be distinct, which would mean that the number of queries would need to be at most 2^l.

Proposition 6 *If $F_k : \{0,1\}^l \to \{0,1\}^L$ is a (t, q, ϵ)-secure PRF then $G_k(x) = (F_k(2x), F_k(2x+1))$, $G_k : \{0,1\}^{l-1} \to \{0,1\}^{2L}$, is a $(t, q/2, 2\epsilon)$-secure PRF.*

Proof: For a random function $R : \{0,1\}^l \to \{0,1\}^L$, $(R(2x), R(2x+1))$ is a random function. Let $D_R = (R(2x), R(2x+1))$ and $D_G = (F_k(2x), F_k(2x+1))$. Suppose there is an adversary A such that $\left|\mathbf{Pr}[A^{D_R} = 1] - \mathbf{Pr}[A^{D_G} = 1]\right| > 2\epsilon$. Then we can build an adversary B to distinguish R from F_k, where B makes two calls to its oracle whenever A needs to use its oracle, i.e. $B^{\mathcal{O}} = A^f$, where $f(x) = (\mathcal{O}(2x), \mathcal{O}(2x+1))$. $\square$

Informally, according to Proposition 6, if we could distinguish between the concatenation of outputs from a PRF F_k and the concatenation of outputs from a random function R, then we could distinguish between F_k and R. Therefore we can safely use the concatenation of outputs from F_k as an XOR mask in S-CTR, and the mode will be secure as long as the counter doesn't wrap.

Finally we consider the case in which the counter is randomly chosen (R-CTR).

Theorem 7 *If $E_k : \{0,1\}^l \to \{0,1\}^L$ is a $(t, \mu/L, \epsilon)$-secure PRF, then R-CTR[E_k] is $(t, q, \mu, 2\epsilon + \mu q/(L \cdot 2^l))$-secure against LOR-CPA (if $\mu < 2^l$).*

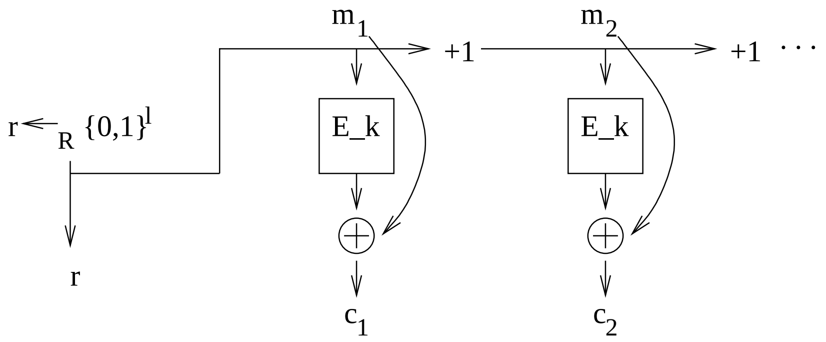

Figure 3: R-CTR

This is a weak result if we use many very short messages. There is no security guarantee if the number of messages is approximately the square root of the size of the counter space. On the other hand, there is decent security if there are just a few long messages.

Lemma 8 *If $R : \{0,1\}^l \to \{0,1\}^L$ is a random function, then R-CTR$[R]$ is $(\infty, q, \mu, \mu q/(L \cdot 2^l))$-secure in the LOR sense against CPA.*

Proof: It is a "bad" event when the counter repeats. Let n_i be the length of message i. Then we formally define a BAD event as follows: BAD is the event where there exists i and j, $i \neq j$, such that $[r_i, \ldots, r_i + n_i) \cap [r_j, \ldots, r_j + n_j) \neq \varnothing$. Equivalently, BAD is the event where there exists $i < j$ such that $r_j \in \{r_i - n_j + 1, \ldots, r_i + n_i - 1\}$. Given that BAD does not happen (i.e. $\overline{\text{BAD}}$), the probability that $A^{E_R \circ \S_i}$ returns 1 is the same for $i = 0$ and $i = 1$. That is,

$$\left|\mathbf{Pr}[A^{\text{R-CTR}[R]\circ\S_0} = 1|\overline{\text{BAD}}] - \mathbf{Pr}[A^{\text{R-CTR}[R]\circ\S_1} = 1|\overline{\text{BAD}}]\right| = 0.$$

Now

$$\begin{aligned}
\mathbf{Pr}[\text{BAD}] &\leq \sum_{i<j} \mathbf{Pr}[r_j \in (r_i - n_j, r_i + n_i)] \\
&\leq \sum_{i<j} \frac{n_i + n_j - 1}{2^l} \\
&\leq \frac{1}{2} \cdot \frac{1}{2^l} \sum_{i,j} (n_i + n_j - 1) \\
&\leq \frac{q}{2^l} \sum n_i \\
&= \frac{q\mu}{2^l \cdot L}
\end{aligned}$$

since $\sum n_i = \mu/L$. □

CTR is an example of a mode suggested 25 years ago which in the past 10 years has been shown to be secure as long as the underlying primitive is secure.