Security of COLM

COLM Team

The COLM proof we provide below follows the proof of the ELmD authenticated encryption design. Here we show the proof for COLM with no intermediate tags. The proof for COLM with intermediate tags follows in the same way the proof of ELmD with intermediate tags and similarly, archives the same security bounds. In both cases, the proof structure and reasoning is preserved.

1 Confidentiality of COLM_{Π}

Theorem 1. *Let $\mathcal{A}$ be an adversary which can make q queries at an aggregrate of total σ associated data and message blocks to distinguish COLM_{Π} with $\mathsf{t} = 0$, from an online cipher chosen uniformly at random. Let $\sigma_{\text{priv}} = \sigma + q$. The online privacy advantage of the adversary $\mathcal{A}$ is given by,*

$$\mathbf{Adv}^{\text{opriv}}_{\mathsf{COLM}_{\Pi}}(\mathcal{A}) \;\leq\; \frac{5\sigma^2_{\text{priv}}}{2^n}.$$

Proof. Let us fix q associate data with initial block, plaintext pairs $P_1 = (A_1, M_1), \ldots, P_q = (A_q, M_q)$ with $|A_i| = a_i, |M_i| = l_i$, $\sigma = \sum_i a_i + l_i$, $\sigma_{\text{priv}} = \sigma + q$. We denote $(P_1, \ldots, P_q)$ by τ_{in}. We assume that all P_i's are distinct.

Step I. Define good online view. A tagged ciphertext tuple $\tau_{out} = (C_1, \ldots, C_q)$ is called **good** online view (belongs to τ_{good}) w.r.t. τ_{in} if (τ_{in}, τ_{out}) is an online view (i.e. $(M_i[1..j] = M_{i'}[1..j]) \Rightarrow (C_i[j] = C_{i'}[j])$) and the following conditions hold:

1. $C_i[j] = C_{i'}[j']$ implies that $j = j'$, $(A_i, M_i[..j]) = (A_{i'}, M_{i'}[..j])$
2. $\forall\ (i, l_i + 1) \neq (i', j'),\ C_i[l_i + 1] \neq C_{i'}[j']$.

One can easily show that,

Lemma 1 (Obtaining a Good view has high probability).

$$\mathsf{Pr}[\tau(\mathcal{A}^{\$_{ol}}) \notin \tau_{good}] \;\leq\; \frac{\sigma^2_{\text{priv}}}{2^n}$$

Step II. High Interpolation Probabilities for good online view. We now fix a good view $\tau = (\tau_{in}, \tau_{out})$ as mentioned above and we will show that,

Lemma 2 (High interpolation probability of COLM). $\forall \tau \in \tau_{good}$,

$$\mathsf{Pr}[\tau(\mathcal{A}^{\mathsf{COLM}_{\Pi}}) = \tau] \geq (1 - \tfrac{4\sigma^2_{\text{priv}}}{2^n}) \times \mathsf{Pr}[\tau(A^{\$^{ol}}) = \tau].$$

Proof. As adversary is deterministic, we restrict to those good views which can be obtained by A. Hence the probability $\mathsf{Pr}[\tau(A^{\mathsf{COLM}}) = \tau]$ is same as

$$\mathsf{Pr}[\mathsf{COLM}_{\Pi}(A_i, M_i) = C_i, 1 \leq i \leq q].$$

Before computing interpolation probability we denote all intermediate variables while computing $\mathsf{COLM}_{\Pi}(A_i, M_i) = C_i$. Let for all i and j whenever defined

1. $AA_i[j] = 3 \cdot L \cdot 2^{j-1} + A_i[j]$, $MM_i[j] = L \cdot 2^{j-1} + M_i[j]$
2. $\Pi(AA_i[j]) = Z_i[j]$, $\Pi(MM_i[j]) = X_i[j]$,
3. $\mathsf{mix}(Z_i, X_i) = Y_i$ and
4. $CC_i[j] = 3^2 \cdot L \cdot 2^{j-1} + C_i[j]$

Now, We call L valid corresponding to a (Z, X) if all the following holds (ensures no undesired collisions in block cipher inputs):

- $SS_i[j] \neq SS_{i'}[j']$ if $j \neq j'$ or $S_i[j] \neq S_{i'}[j]$ for $S \in \{A, M, C\}$,
- $MM_i[j] \neq AA_{i'}[j']$,
- $Z_i[j] \neq CC_{i'}[j']$,
- $X_i[j] \neq CC_{i'}[j']$.

Remark 1. Notice, that for ELmD, instead of the last two inequalities, we needed the inequalities (a) $AA_i[j] \neq CC_{i'}[j']$ and (b) $MM_i[j] \neq CC_{i'}[j']$. This is the only technical difference in the proof.

Now, the primitivity of 2 and uniform independent choice of L ensures that each equality violating has probability 2^{-n} and there are at most $2.\sigma^2_{\text{priv}}$ equalities possible. So, using union bound we have

$$\mathsf{Pr}[L \text{ is valid}] \geq (1 - \tfrac{2\sigma^2_{\text{priv}}}{2^n}).$$

Now, for any fix L, applying the consistent collision relations for linear functions, the conditional interpolation probability is

$$\sum_{(Z,X)} \frac{\#\Pi : \Pi(MM) = X, \Pi(AA) = Z, \Pi(Y) = CC}{\#\Pi}$$

$$\geq (1 - \frac{2\sigma^2_{\text{priv}}}{2^n}).2^{-nP}.$$

This proof is identical to the one used for ELmD. Now, multiplying the above probability for validness of L the proof of the lemma completes.

2 Integrity of COLM_{Π}

Theorem 2. *Let $\mathcal{A}$ be an adversary which can make q forward queries and tries to forge s many times at an overall aggregrate of total σ associated data and message blocks. Let $\sigma_{\text{auth}} = \sigma + q$. Then the forging advantage of the adversary is given by,*

$$\mathbf{Adv}^{\text{auth}}_{\mathit{COLM}_{\Pi}}(\mathcal{A}) \;\leq\; \frac{9\sigma^2_{\text{auth}}}{2^n} + \frac{s}{2^n}.$$

Proof. We know that PRF implies MAC. We use similar concept to bound authenticity: for any forgery $\mathcal{B}$, there is a distinguisher $\mathcal{A}$ such that

$$\mathbf{Adv}^{\text{auth}}_{F}(\mathcal{B}) \leq \mathbf{Adv}^{\mathcal{O},\$}_{(F,T)}(\mathcal{A}) + \frac{s}{2^n} \tag{1}$$

Now we will bound $\mathbf{Adv}^{\$_{ol},\$}_{(F,T)}(\mathcal{A})$ for COLM.

A (F,T)-view of a distinguisher A is the pair $v = (\tau_F, \tau_T)$ where $\tau_F = (A_i, M_i, C_i)_{1\leq i\leq q}$ is an q-tuple of F-online view and $\tau_T = (D_j, C_j)_{q<j\leq q+s}$ is an s-tuple non-trivial T-view. Suppose, $|A_i| = a_i$, $|M_i| = l_i$ and $|C_i| = l_i + 1$. Let $\sigma = \sum_{i=1}^{q+s}(a_i + l_i)$ and $\sigma_{auth} = \sigma + q$ (the total number of tagged ciphertext blocks).

Step I. Define good online view. A (F,T) view is called **good** online forge view if both the following holds:

1. τ_F is **good** online forge view (as defined in the privacy prove).
2. $\forall\ q < j \leq q + s$, $C_j[l_j + 1]$'s are fresh - distinct and different from all other $C_i[j]$'s.

Lemma 3 (Realizing Good Forge View has high probability). *For all adversary $\mathcal{A}$,*

$$\mathsf{Pr}[\tau(\mathcal{A}^{\$_{ol},\$}) \notin \tau_{good}] \leq \frac{(q + \sum_{i=1}^{q} l_i)^2}{2^{n+1}} + \frac{s(q + s + \sum_{i=1}^{q+s} a_i + l_i)}{2^n} \leq \frac{2\sigma^2_{auth}}{2^n}$$

Step II. High Interpolation Probabilities for good online view. We now fix a good view $\tau = (\tau_{in}, \tau_{out})$ as mentioned above and we will show that,

Lemma 4 (Good Forge View has high interpolation probability). *For any good (F,T)-view τ, we have*

$$\mathsf{Pr}[F(A_i, M_i) = C_i, \forall i \leq q,\ T(A_j, C_j[..l_j]) = C_j[l_j{+}1], q < j \leq q{+}s] \ \geq\ \frac{(1 - 7\sigma^2_{\text{auth}}/2^n)}{2^{n(P+s)}}$$

Proof. We choose $X_1, \ldots, X_q$ and then $Y_{q+1}, \ldots, Y_{s+q}$ which fix all internal X and Y values except the last block for the s many T-queries. We choose valid L which fixes MM's for the first q messages and, CC's and AA's for all $s + q$ queries. We can then choose MM for these s queries so that checksums are all fresh and for all these fresh checksums we can ensure last Y blocks fresh by choosing X blocks appropriately. Now we make these choices:

Choices of Valid L. We first define valid L-triples as defined in privacy. L is called valid w.r.t. the fixed good (F,T)-view τ if the computed MM, AA, CC values satisfy the following (ensures no undesired collisions in block cipher inputs):

- $SS_i[j] \neq SS_{i'}[j']$ if $j \neq j'$ or $S_i[j] \neq S_{i'}[j]$ for $S \in \{A, M, C\}$,
- $MM_i[j] \neq AA_{i'}[j']$,
- $MM_i[l_i] \neq MM_j[l_j]$ if C_j, $j > q$, is a strictly prefix of C_i, $i \leq q$ and $A_i = A_j$,

- $Z_i[j] \neq CC_{i'}[j']$,
- $X_i[j] \neq CC_{i'}[j']$.

Remark 2. For ELmD, instead of the last two inequalities, we needed the inequalities (a) $AA_i[j] \neq CC_{i'}[j']$ and (b) $MM_i[j] \neq CC_{i'}[j']$. This is the only technical difference in the proof.

The simple counting argument with union bound applied to all individual bad events proves the following result.

$$\mathsf{Pr}[L \text{ is valid}] \ \geq \ (1 - \tfrac{2\sigma_{\text{auth}}^2}{2^n}).$$

Now one can (i) first choose valid Z, X, Y except except the last blocks for the last s queries, (ii) given the choices of valid X, Y, Z one can choose all those MM values for which $X_j[i]$'s are fresh, (iii) finally for any such previous choices, one can choose the blocks of $X_j[l_j + 1]$, $j > q$ so that the last block of Y_j's are fresh.

Now, for any fix L, applying the consistent collision relations for linear functions, the conditional interpolation probability is

$$\sum_{(Z,X)} \frac{\#\Pi : \Pi(MM) = X, \Pi(AA) = Z, \Pi(Y) = CC}{\#\Pi}$$

$$\geq (1 - \frac{7\sigma_{\text{auth}}^2}{2^n}).2^{-n(P+s)}.$$

This proof is identical to the one used in ELmD. Now,multiplying the above probability for validness of L the proof of the lemma completes.