ETH zürich

Departement of Computer Science

SCB Mode: Semantically Secure Length-Preserving Encryption

Fabio Banfi

FSE 2023, 24 March 2023

Outline

1. Background and Motivation
2. Length-Preserving Encryption/Enciphering (LPE)
3. SCB Mode of Encryption: Semantically Secure LPE
4. Conclusions

Outline

Modes of Operation

Turn block cipher $\mathfrak{B} = (E, D)$ into encryption scheme $\Pi = (\mathcal{E}, \mathcal{D})$:

For $M = M_1 \| \cdots \| M_\ell \in \{0,1\}^{\ell n}$: compute ciphertext $C \in \{0,1\}^{\ell n + \lambda}$ (with ***expansion factor*** λ)

- An ***insecure*** way: Electronic Codebook (ECB) Mode ($\boldsymbol{\lambda = 0}$):

$$\mathcal{E}_K(M) \doteq E_K(M_1) \| \cdots \| E_K(M_\ell) \in \{0,1\}^{\ell n}$$

- A ***secure*** way: Cipher Block Chaining (CBC) Mode ($\boldsymbol{\lambda = n}$): Sample $R \xleftarrow{\$} \{0,1\}^n$, then

$$\mathcal{E}_K(M) \doteq R \| \underbrace{E_K(R \oplus M_1)}_{C_1} \| \underbrace{E_K(C_1 \oplus M_2)}_{C_2} \| \cdots \| E_K(C_{\ell-1} \oplus M_\ell) \in \{0,1\}^{\ell n + n}$$

Both can be adapted to handle *any* $M \in \{0,1\}^{\geq n}$ via ciphertext stealing (CTS)

Motivation

Question: *Can we get the best of both (**secure** and $\lambda = 0$)?*

What if we have many *short* messages to be transmitted, and communication is expensive? E.g.:

- Each day m messages need to be transmitted
- Each message consists of b blocks (defined by the underlying block cipher)

Conventional IND-CPA scheme: $c_0 \doteq m(b+1)$ transmitted blocks

Encryption without expansion: $c_1 \doteq mb$ transmitted blocks

$\Longrightarrow$ If b small and m large: $c_0 \approx 2 \cdot c_1$!

Can we avoid expansion while retaining semantic security? Seems impossible, but let's see ...

Outline

1. Background and Motivation
2. Length-Preserving Encryption/Enciphering (LPE)
3. SCB Mode of Encryption: Semantically Secure LPE
4. Conclusions

Encryption Schemes with $\lambda = 0$

With $\lambda = 0$, Π *cannot* be semantically secure, why? If $\lambda = 0$, then $\mathcal{E}_K(\cdot)$ must be deterministic!

Therefore, for any $K \in \mathcal{K}$ and any $t \in \mathbb{N}$, algorithm $\mathcal{E}_K$ is a *permutation* on $\{0,1\}^t$

Known as Length-Preserving Encryption (LPE), but should be called: Length-Preserving *Enciphering*!

Alternatively, Π can be seen as a variable-input-length (VIL) *block cipher*

Back to our question: Can we design a semantically secure encryption scheme with $\lambda = 0$?

Yes! If we relax **correctness** to *not* be perfect but only *computational!*

Therefore, $\mathcal{E}_K$ might *not be* a permutation on $\{0,1\}^t$, and should be **stateful**

Semantically Secure Length-Preserving *Encryption*?

Definition (Length-Preserving *Stateful* Encryption (LPSE))

A pair Π of algorithms:

- $\mathcal{E} : \mathcal{K} \times \{0,1\}^{\geq n} \times \mathcal{S} \to \{0,1\}^{\geq n} \times \mathcal{S}$,
- $\mathcal{D} : \mathcal{K} \times \{0,1\}^{\geq n} \times \mathcal{T} \to \{0,1\}^{\geq n} \times \mathcal{T}$,

s.t. for any $K \in \mathcal{K}$, encryption state $\mathbf{S} \in \mathcal{S}$, and decryption state $\mathbf{T} \in \mathcal{T}$:

- $\mathcal{E}(K, \cdot\,; \mathbf{S})$ and $\mathcal{D}(K, \cdot\,; \mathbf{T})$ are *efficiently computable*
- For any $t \in \mathbb{N}$, and $M, C \in \{0,1\}^t$:
  - $\mathcal{E}(K, M; \mathbf{S}) \in \{0,1\}^{|M|} \times \mathcal{S}$ $\quad [C \leftarrow \mathcal{E}_K^{\mathbf{S}}(M)$ denotes $(C, \mathbf{S}') \leftarrow \mathcal{E}(K, M; \mathbf{S});\ \mathbf{S} \leftarrow \mathbf{S}']$
  - $\mathcal{D}(K, C; \mathbf{T}) \in \{0,1\}^{|C|} \times \mathcal{T}$ $\quad [M \leftarrow \mathcal{D}_K^{\mathbf{T}}(C)$ denotes $(M, \mathbf{T}') \leftarrow \mathcal{D}(K, C; \mathbf{T});\ \mathbf{T} \leftarrow \mathbf{T}']$

Note: There is no correctness requirement in the definition!

LPSE: Security and Correctness

Let $[\,] \in \mathcal{S}, \mathcal{T}$ denote the initial empty encryption/decryption state

Definition (LPSE Semantic Security)

$\Pi = (\mathcal{E}, \mathcal{D})$ is a *semantically secure LPSE scheme* if for any IND-CPA adversary A, its advantage

$$\mathbf{Adv}^{\text{ind-cpa}}_{\Pi}(A) \doteq \Pr\left[A^{\mathcal{E}^{\mathbf{S}}_K(\cdot)} \Rightarrow 0 \,\middle|\, K \xleftarrow{\$} \mathcal{K}, \mathbf{S} \leftarrow [\,]\right] - \Pr\left[A^{\$^{|\cdot|}} \Rightarrow 0\right]$$

is negligible

Definition (LPSE Correctness)

$\Pi = (\mathcal{E}, \mathcal{D})$ is a *correct LPSE scheme* if for any COR adversary A, its advantage

$$\mathbf{Adv}^{\text{cor}}_{\Pi}(A) \doteq \Pr\left[A^{\mathcal{D}^{\mathbf{T}}_K \circ \mathcal{E}^{\mathbf{S}}_K(\cdot)} \Rightarrow 0 \,\middle|\, K \xleftarrow{\$} \mathcal{K},\ \mathbf{S}, \mathbf{T} \leftarrow [\,]\right] - \Pr\left[A^{\mathsf{id}(\cdot)} \Rightarrow 0\right]$$

is negligible

Outline

1. Background and Motivation
2. Length-Preserving Encryption/Enciphering (LPE)
3. SCB Mode of Encryption: Semantically Secure LPE
4. Conclusions

SCB: The Idea

We introduce a new mode of operation that turns a block cipher $\mathfrak{B} = (E, D)$ into an LPSE $\Pi = (\mathcal{E}, \mathcal{D})$

Secure Codebook (SCB): Can be interpreted as a secure variant/patch of ECB

Observation: ECB insecure as soon as a block $\hat{M} \in \{0,1\}^n$ is repeated *within* or *across* plaintexts

$\Longrightarrow$ Use state to keep track of blocks seen so far, and *on repeated blocks do something different!*

But what to do exactly? We need to signal to the receiver that this block is a repetition of $\hat{M}$

This inevitably would introduce errors, since a subspace of $\{0,1\}^n$ must represent such signals!

But we can be clever about the choice of such subspace :)

SCB: Encryption

Idea: Let σ and τ be such that $\sigma + \tau \leq n$, $K_1 \in \{0,1\}^\kappa$ (for $\mathfrak{B}$), and $K_2 \in \{0,1\}^n$ (pad), and consider:

- A compression function $H : \{0,1\}^n \to \{0,1\}^\tau$
- A look-up table $\mathbf{S} : \{0,1\}^\tau \to \{0,1\}^\sigma$ (for $h \in \{0,1\}^\tau$, $\mathbf{S}[h] \in \{0,1\}^\sigma \cup \{\perp\}$)

Then for each block M_i:

1. Get $h \leftarrow H(M_i)$, and check whether h is in $\mathbf{S}$, i.e., $\mathbf{S}[h] \neq \perp$ (approximates "M_i is a repetition")
2. If **not** (M_i is a *new* block), then compute $C_i \leftarrow \mathfrak{B}.E_{K_1}(M_i)$ (plain ECB) and set $\mathbf{S}[h] \leftarrow 0^\sigma$
3. If **yes** (M_i is *probably a repeated* block, but **might be wrong**), then:
   - Let $R \leftarrow (0^{n-\sigma-\tau} \,\|\, \mathbf{S}[h] \,\|\, h) \in \{0,1\}^n$, and compute $C_i \leftarrow \mathfrak{B}.E_{K_1}(K_2 \oplus R)$
   - Set $\mathbf{S}[h] \leftarrow (\mathbf{S}[h] + 1) \bmod 2^\sigma$

SCB: Decryption

But how do we decrypt now?

We need to distinguish between normal blocks and repetition signals!

Let $\sigma, \tau, K_1, K_2, H$ as before, and consider look-up table $\mathbf{T} : \{0,1\}^\tau \to \{0,1\}^n$ (approximates "H^{-1}")

Then for each block C_i:

1. Get $M_i \leftarrow \mathfrak{B}.D_{K_1}(C_i)$ (plain ECB)
2. Compute $R \leftarrow K_2 \oplus M_i$ and $h \leftarrow R \bmod 2^\tau$, and check whether $R < 2^{\sigma+\tau}$ and $\mathbf{T}[h] \neq \perp$
3. If **not** (C_i is a *not* a repetition signal), then keep M_i and set $\mathbf{T}[H(M_i)] \leftarrow M_i$
4. If **yes** (C_i is *probably a repetition signal*, but **might be wrong**), then set $M_i \leftarrow \mathbf{T}[h]$

SCB: Security and Correctness

We show that SCB is secure if the underlying block cipher $\mathfrak{B} = (E, D)$ is a secure PRP

Theorem (Security)

For any $\mathrm{IND\text{-}CPA}$ *adversary* A *querying* $\beta \leq 2^{\sigma}$ *blocks we can construct a* PRP *adversary* B *such that*

$$\mathbf{Adv}^{\text{ind-cpa}}_{\mathsf{SCB}[\mathfrak{B},H]}(A) \leq \mathbf{Adv}^{\text{prp}}_{\mathfrak{B}}(B) + \frac{\beta^2}{2^n}$$

We show that SCB is correct if the underlying compression function H is collision resistant

Theorem (Correctness)

For any COR *adversary* A *querying* β *blocks we can construct a* CR *adversary* B *such that*

$$\mathbf{Adv}^{\text{cor}}_{\mathsf{SCB}[\mathfrak{B},H]}(A) \leq \mathbf{Adv}^{\text{cr}}_{H}(B) + \frac{2^{\sigma}\beta^2}{2^n}$$

Outline

1. Background and Motivation
2. Length-Preserving Encryption/Enciphering (LPE)
3. SCB Mode of Encryption: Semantically Secure LPE
4. Conclusions

Conclusions

We introduced the first IND-CPA-secure length-preserving encryption scheme (for *any* length via CTS)

In the paper we also consider a variant that is secure and correct even if ciphertexts are *reordered*

We also identify possible improvements for future work:

- Checking counters upon decryption to remove factor 2^{σ} in correctness
- Is it possible to have better *state size growth*? (probably can't be zero)
- Are there other schemes with better security/correctness bounds?

Thank you for your attention!