I240 暗号理論 2019
共通鍵暗号とメッセージ認証 (1)

2019/10/16（修正 10/17） 講師 藤崎

1 はじめに

情報理論的に安全な共通鍵暗号、メッセージ認証コード (MAC) について。

2 共通鍵暗号

2.1 共通鍵暗号 (Symmetric-Key Encryption) の定義

共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ とは、二つの集合と二つのアルゴリズムの組であり，次のように定義される：

- 鍵生成空間 $\mathcal{K}$: 秘密鍵 k の集合.
- 平文空間 $\mathcal{M}$: 平文 m の集合.
- 鍵生成: $\mathcal{K}$ から一様ランダムに秘密鍵 k を選ぶ。この試行を $k \leftarrow \mathcal{K}$ と書く。
- 暗号化アルゴリズム $\mathbf{E}$: 秘密鍵 $k \in \mathcal{K}$, $m \in \mathcal{M}$ を入力としてとり，暗号文 c を出力するアルゴリズム。この試行を $c \leftarrow \mathbf{E}_k(m)$ と書く. 出力は確率的 (probablistic) でもよい $c = \mathbf{E}_k(m; r)$（r は $\mathbf{E}$ に入力される乱数を表す)。
- 復号アルゴリズム $\mathbf{D}$: 秘密鍵 s と暗号文 c を入力としてとり, 平文 m を出力する確定的 (deterministic) アルゴリズム。この試行を $m \leftarrow \mathbf{D}_k(c)$ と書く。

さらに、$\mathbf{D}$ は全ての鍵 $k \in \mathcal{K}$, 全ての平文 $m \in \mathcal{M}$ に対して、常に $\mathbf{D}_k(\mathbf{E}_k(m)) = m$ を満たす。この条件を共通鍵暗号の Correctness 条件という。すなわち、

$$\text{Correctness:} \quad \forall m \in \mathcal{M} \quad \Pr_{k \leftarrow \mathcal{K}}[\mathbf{D}_k(\mathbf{E}_k(m)) = m] = 1.$$

2.2 共通鍵暗号の完全秘匿性 (Perfect Secrecy)

共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たすとは、平文に対する情報量が、暗号文 c を見る前と後で全く変わらないときを言う。

定義 1 (Shannon [Sha49]) 共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性 (perfect secrecy) を満たすとは、$\mathcal{M}$ 上定義されるの任意の確率分布 X, 任意の平文 $m \in \mathcal{M}$, 任意の暗号文 c に対して、

$$\Pr_{X,K}[X = m \wedge \mathbf{E}_K(X) = c] = \Pr_X[X = m] \cdot \Pr_{X,K}[\mathbf{E}_K(X) = c].$$

が成り立つときを言う。ここで、K は $\mathcal{K}$ から k を一様ランダムに選ぶ分布に従う確率変数とする。

すなわち、$\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性とは、任意の確率変数 X と $\mathbf{E}_K(X)$ が（平文の分布と暗号文の分布が）独立であること。

SKE : perfectly secret $\overset{\text{def}}{\iff}$

$$\forall X \, \forall m \, \forall c : \Pr_{X,K}[X = m \wedge \mathbf{E}_K(X) = c] = \Pr_X[X = m] \cdot \Pr_{X,K}[\mathbf{E}_K(X) = c].$$

命題 2　以下は全て等しい。

1. 共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たす。
2. $\mathcal{M}$ 上定義されるの任意の確率分布 X, 任意の平文 $m \in \mathcal{M}$, 任意の暗号文 $c \in \mathsf{supp}(\mathbf{E}_K(X)) := \{c \mid \Pr_{X,K}[\mathbf{E}_K(X) = c] > 0\}$ に対して、

$$\Pr_{X,K}[X = m \mid \mathbf{E}_K(X) = c] = \Pr_X[X = m].$$

3. 任意の平文 $m_0, m_1 \in \mathcal{M}$ に対して、確率変数 $\mathbf{E}_K(m_0)$ と $\mathbf{E}_K(m_1)$ が同一の確率分布に従う。

$$\mathbf{E}_K(m_0) \equiv \mathbf{E}_K(m_1).$$

4. $\mathcal{M}$ 上定義されるの任意の確率分布 X に対して、

$$H(X) = H(X \mid \mathbf{E}_K(X))$$

が成り立つ。ここで $H(X)$ は確率変数 X の Shannon entropy を表す（後述)。

5. $\mathcal{M}$ 上定義されるの任意の確率分布 X に対して、

$$H(X, Y) = H(X) + H(\mathbf{E}_K(X)).$$

問題 3　上記命題を証明せよ。

2.3 One-Time Pad (Vernam Cipher)

共通鍵暗号の定義に基づいて記述すると次のようになる。

- 鍵空間 $\mathcal{K} = \{0,1\}^\kappa$.
- 鍵生成: $\mathcal{K}(= \{0,1\}^\kappa)$ から一様ランダムに秘密鍵 k を選ぶ。
- 暗号化アルゴリズム $\mathbf{E}$. 秘密鍵 k と平文 $m \in \mathcal{M} := \{0,1\}^\kappa$ を受け取り、暗号文 $c = \mathbf{E}_k(m) := m \oplus k$ を出力する。
- 復号アルゴリズム $\mathbf{D}$. 秘密鍵 k と暗号文 c を受け取り、平文 $m = \mathbf{D}_k(c) := c \oplus k$ を出力する.

問題 4　OTP は、Correctness を満たすことを確認せよ。

問題 5　OTP は、完全秘匿性を満たすことを示せ。

問題 6（additive OTP）　n を任意の正の整数とし、$\mathcal{K} = \mathcal{M} = \mathbb{Z}/n\mathbb{Z}$, $\mathbf{E}_k(m) := m + k \bmod n$, $\mathbf{D}_k(c) := c - k \bmod n$ とすると、この additive OTP も完全秘匿性を満たすことを示せ。

問題 7（multiplicative OTP）　n を任意の正の整数とし、$\mathcal{K} = (\mathbb{Z}/n\mathbb{Z})^\times$, $\mathcal{M} = \mathbb{Z}/n\mathbb{Z}$, $\mathbf{E}_k(m) := m \cdot k \bmod n$, $\mathbf{D}_k(c) := c \cdot k^{-1} \bmod n$ とする。この方式は完全秘匿性を満たすだろうか？この方式で、$\mathcal{M} = (\mathbb{Z}/n\mathbb{Z})^\times$ とした場合はどうなるか？

2.4 完全秘匿性を満たす必要条件

定理 8（Shannon [Sha49])　共通鍵暗号 $\mathsf{SKE} = (\mathcal{K}, \mathcal{M}, \mathbf{E}, \mathbf{D})$ が完全秘匿性を満たすならば、$|\mathcal{K}| \geq |\mathcal{M}|$ が成り立つ。

$$\mathsf{SKE}: \text{ perfectly secret } \Longrightarrow |\mathcal{K}| \geq |\mathcal{M}|$$

暗号化アルゴリズム $\mathbf{E}$ が確定的 (deterministic) でない、つまり確率的 (probablistic) でも成り立つ。

(証明) SKE が、$|\mathcal{K}| < |\mathcal{M}|$ にも関わらず、完全秘匿性を満たしたとすると矛盾であることを証明する。

$m_0 \in \mathcal{M}$, $k_0 \in \mathcal{K}$ を選び、$c \leftarrow \mathbf{E}_{k_0}(m_0)$ を計算する。$S(c) := \{\mathbf{D}_k(c) \,|\, k \in \mathcal{K}\}$ と定義する。すると、復号アルゴリズムの確定性から $S(c) \leq |\mathcal{K}|$. 一方、仮定より $|\mathcal{K}| < |\mathcal{M}|$ であるので、

$$S(c) < |\mathcal{M}|$$

が成り立つ。よって、$\mathcal{M} \backslash S(c) \neq \emptyset$ であるから、$m \in \mathcal{M} \backslash S(c)$ が選べる。その選び方より、全ての $k \in \mathcal{K}$ に対して、$\mathbf{E}_k(m) \neq c$. よって、($\Pr[X = m_0] > 0$ なる) 全ての確率変数 X に対して、

$$\Pr_{X,K}[X = m \,|\, E_K(X) = c] = 0. \tag{1}$$

いま、X を $\mathcal{M}$ 上の一様分布に従う確率変数とする。すると、

$$\Pr_X[X = m] = \frac{1}{|\mathcal{M}|}. \tag{2}$$

ここで、$c \in \mathsf{supp}(E_K(X))$ であり、SKE は完全秘匿性を満たすので、式 (1) と式 (2) の値は一致しなければならないが、一致しないので矛盾。(よって、対偶により、) SKE が完全秘匿性を満たすなら、$|\mathcal{K}| \geq |\mathcal{M}|$ となる。 ■

2.5 Shannon Entropy

確率変数 X に対して $X = x$ なる事象の起こりにくさの尺度を示す値として情報量 (information content) がある。確率変数 X における $X = x$ を表す情報量は $\mathsf{Info}(x) = -\log_2(p(x))$ で定義される ($p(x) := \Pr[X = x]$). $p(x, y) := \Pr[X = x \wedge Y = y]$ と $p(x|y) := \Pr[X = x \,|\, Y = y]$ とし、$\mathsf{Info}(x, y) := -\log_2(p(x, y))$, $\mathsf{Info}(x|y) := -\log_2(p(x|y))$ と定義する。その時、次が成立することは容易にわかる。$\mathsf{Info}(x, y) = \mathsf{Info}(x) + \mathsf{Info}(y|x) = \mathsf{Info}(y) + \mathsf{Info}(x|Y)$. X, Y が独立ならば、$\mathsf{Info}(x, y) = \mathsf{Info}(x) + \mathsf{Info}(y)$.

確率変数 X に対する Shannon Entropy (または平均情報量) を、

$$H(X) := \sum_{x \in \mathcal{X}} p(x) \cdot \mathsf{Info}(x)$$

で定義する。

確率変数 (X, Y) の (X と Y の結合) Shannon Entropy は、

$$H(X, Y) := \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \cdot \mathsf{Info}(x, y)$$

で定義される。一方、条件付き Shannon Entropy は

$$H(X|Y) := \sum_{y \in \mathcal{Y}} H(X|Y = y)$$

で定義される。$H(X|Y = y) = \sum_{x \in \mathcal{X}} p(x|y) \cdot \mathsf{Info}(x|y)$ より、

$$\begin{aligned} H(X|Y) &= \sum_{y \in \mathcal{Y}} p(y) \sum_{x \in \mathcal{X}} p(x|y) \cdot \mathsf{Info}(x|y) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \cdot \mathsf{Info}(x|y). \end{aligned}$$

命題 9（Chain Rule）

$$H(X,Y) = H(X) + H(Y|X) = H(Y) + H(X|Y).$$

(証明) 式変形から明らか。

$$\begin{aligned}
H(X,Y) &:= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x,y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot (\mathsf{Info}(x) + \mathsf{Info}(x|y)) \qquad (\text{by } \mathsf{Info}(x,y) = \mathsf{Info}(x) + \mathsf{Info}(x|y)) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x|y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + H(X|Y) \qquad (\text{by } H(X|Y) = \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x|y)) \\
&= \sum_{x\in\mathcal{X}} p(x)\cdot \mathsf{Info}(x) + H(X|Y) \qquad (\text{by } p(x) = \sum_{y\in\mathcal{Y}} p(x,y)) \\
&= H(X) + H(X|Y).
\end{aligned}$$

命題 10 X, Y が独立な確率変数であれば、

$$H(X,Y) = H(X) + H(Y).$$

(証明) X, Y が独立なら $\mathsf{Info}(x,y) = \mathsf{Info}(x) + \mathsf{Info}(y)$. さらに、$p(x) = \sum_{y\in\mathcal{Y}} p(x,y)$ に注意すれば式変形から明らか。

$$\begin{aligned}
H(X,Y) &:= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x,y) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot (\mathsf{Info}(x) + \mathsf{Info}(y)) \\
&= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(x) + \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}} p(x,y)\cdot \mathsf{Info}(y) \\
&= H(X) + H(Y)
\end{aligned}$$

命題 11 以下は全て等しい。

1. X と Y が独立な確率変数
2. 全ての $x \in \mathcal{X}$, $y \in \mathcal{Y}$ に対して、
$$\Pr[X = x \wedge Y = y] = \Pr[X = x]\cdot\Pr[Y = y].$$
3. 全ての $x \in \mathcal{X}$, $y \in \mathcal{Y}$ に対して、
$$\Pr[X = x] = \Pr[X = x|Y = y].$$
4. X, Y の結合エントロピーが各エントロピーの和
$$H(X,Y) = H(X) + H(Y).$$
5. X の Shannon entropy が、Y が起きた後にも減少しない
$$H(X) = H(X|Y).$$

問題 12 命題 11 を証明せよ。

3 メッセージ認証コード (Message Authentication Code)

3.1 メッセージ認証コード (MAC) の定義

メッセージ認証 $\mathsf{MAC} = (\mathcal{K}, \mathcal{M}, \mathsf{S}, \mathsf{V})$ とは、二つの集合と二つのアルゴリズムの組であり、次のように定義される。

- 鍵生成空間 $\mathcal{K}$: 秘密鍵 k の集合。
- 平文空間 $\mathcal{M}$: 平文 m の集合。
- 鍵生成: $\mathcal{K}$ から一様ランダムに秘密鍵 k を選ぶ。この試行を $k \leftarrow \mathcal{K}$ と書く。
- MAC 生成アルゴリズム S: 秘密鍵 $k \in \mathcal{K}$, 平文 $m \in \mathcal{M}$ を入力としてとり、メッセージ認証子 τ を出力するアルゴリズム。この試行を $\tau \leftarrow \mathsf{S}(k, m)$ と書く。出力は確率的でもよい。
- MAC 検証アリゴリズム V: 秘密鍵 $k \in \mathcal{K}$, 平文 $m \in \mathcal{M}$, 認証子 τ を入力としてとり、認証子の正しさを判定するアルゴリズム（正しいと判断した場合は、1 を、それ以外の場合は 0 を出力するものとする)。この試行を $b \leftarrow \mathsf{V}(k, m, \tau)$ と書く $(b \in \{0, 1\})$。

さらに、V は、全ての鍵 $k \in \mathcal{K}$, 全ての平文 $m \in \mathcal{M}$, 全ての正当な認証子 $\tau \in \mathsf{S}(k, m)$ に対して、常に $\mathsf{V}(k, m, \tau) = 1$ を満たす。この条件をメッセージ認証の Correctness 条件という。すなわち、

$$\text{Correctness:} \quad \forall m \in \mathcal{M} \quad \Pr[k \leftarrow \mathcal{K}; \tau \leftarrow \mathsf{S}(k, m) : \mathsf{V}(k, m, \tau) = 1] = 1.$$

3.2 MAC の安全性

メッセージ認証 $\mathsf{MAC} = (\mathcal{K}, \mathcal{M}, \mathsf{S}, \mathsf{V})$ の安全性は、次のようなゲームを通じて定義される。

■EUF-CMA ゲーム（MAC 版） 敵 (adversary) A とチャレンジャー C の間で行われるメッセージ認証方式 MAC の安全性を試すゲームである [*1]。

1. C は鍵空間 $\mathcal{K}$ から一様ランダムに鍵 k を選ぶ。$k \leftarrow \mathcal{K}$.
2. A は平文 m を C に送り、対応する認証子 $\tau \leftarrow \mathsf{S}(k, m)$ を答えとしてもらう。平文 m の選び方は A の戦略による。A は合計 q 回まで質問 (query) を送ることができる。A が、平文を質問しその答えとして認証子をもらう行為を、A の $\mathsf{S}(k, \cdot)$ オラクルへのアクセスと呼ぶ。$\mathsf{S}(k, \cdot)$ オラクルへのアクセス履歴を $L = \{(m_1, \tau_1), \ldots, (m_q, \tau_q)\}$ とし、アクセス平文の集合を特に $L(m) = \{m_1, \ldots, m_q\}$ と書くことにする。
3. A は、平文と認証子の組 (m^*, τ^*) を出力する。
4. もし、m^* が、$\mathsf{S}(k, \cdot)$ オラクルへ質問したことのない新しい平文であり $(m^* \notin L(m))$, $\mathsf{V}(k, m^*, \tau^*) = 1$ を満たすのであれば、A の勝ちと定義する。

A が、メッセージ認証方式 MAC に対する EUF-CMA ゲームに勝つ確率を、A の MAC に対するアドバンテージと呼び、

$$\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{A,\mathsf{MAC}}(q, n) := \Pr[k \leftarrow \mathcal{K};\ (m^*, \tau^*) \leftarrow A^{\mathsf{S}(k,\cdot)} : \mathsf{V}(k, m^*, \tau^*) = 1 \text{ and } m^* \notin L(m)].$$

ここで、n は、秘密鍵のサイズ、すなわち $n = \log |\mathcal{K}|$.

[*1] チェレンジャーは明示的には表現されないときもある。

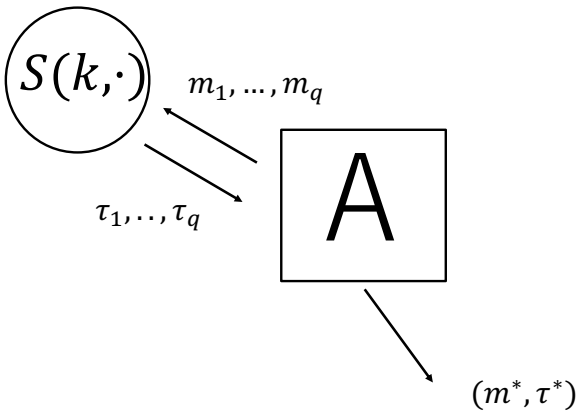

図 1　メッセージ認証の EUF-CMA ゲーム

定義 13（メッセージ認証の安全性） 任意の（計算量やメモリー量を制限しない）敵 A に対して、$\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{A,\mathsf{MAC}}(q,n) = O(2^{-n})$ であるとき、メッセージ認証方式 MAC は q 回までの質問に対して（情報理論的に）安全と呼ぶ。

3.3　One-Time MAC (OT-MAC)

メッセージ認証方式の定義に基づいて記述すると以下のようになる。

- 鍵生成空間 $\mathcal{K} = \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. ただし、$p$ は素数。
- 平文空間 $\mathcal{M} = \mathbb{Z}/p\mathbb{Z}$.
- 鍵生成: $(\alpha, \beta) \leftarrow \mathcal{K}$.
- MAC 生成アルゴリズム S: 秘密鍵 $(\alpha, \beta) \in \mathcal{K}$, 平文 $m \in \mathcal{M}$ を入力としてとり、メッセージ認証子 $\tau = (\alpha m + \beta) \bmod p$ を出力するアルゴリズム。
- MAC 検証アリゴリズム V: 秘密鍵 $(\alpha, \beta) \in \mathcal{K}$, 平文 $m \in \mathcal{M}$, 認証子 τ を入力としてとり、$\tau = (\alpha m + \beta) \pmod p$ ならば、1 を出力、それ以外の場合は 0 を出力する。

定理 14　OT-MAC は、任意の敵に対して、$\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{A,\mathsf{otMAC}}(1,n) = \frac{1}{p}$. ただし、$n = \log(p)$.

OT-MAC は、one-time 安全であるともいう。

(証明) A がオラクルアクセスをして得た平文と認証子の組を (m, τ) とする。すると、$\tau = \alpha m + \beta \pmod p$ を満たしている。(α, β) は鍵生成時に一様に選んでいるから、直線 $\tau = \alpha m + \beta \pmod p$ 上、一様に分布している。A が (m^*, τ^*) を出力するということは、$m^* \neq m \pmod p$ であるから、傾きの違う直線 $\tau^* = \alpha m^* + \beta \pmod p$ を決定することであり、その交点により、(α, β) を決定することを意味する。つまり、A が (m, τ) を見た時点では、(α, β) は $\tau = \alpha m + \beta \pmod p$ 上、一様に分布するという情報しかない。その状態で、(α, β) を正しく推測できる確率は p^{-1} しかないので、A がゲームに勝てる確率は高々 p^{-1} しかない。よって、$\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{A,\mathsf{otMAC}}(1,n) = \frac{1}{p}$.　■

問題 15　任意の q に対して、$\mathsf{Adv}^{\mathsf{euf\text{-}cma}}_{A,\mathsf{MAC}}(q,n) = O(2^{-n})$ となるメッセージ認証方式を考えよ。

参考文献

[Sha49] Claude E. Shannon. Communication theory of secrecy systems. *Bell Systems Technical Journal*, 28(4):656–715, 1949.