［入門講演］

情報理論的安全性

～さまざまな視点から～

岩本 貢

電気通信大学 大学院情報理工学研究科 情報学専攻

2017 年 9 月 6 日

第 6 回誤り訂正符号のワークショップ@山口湯田温泉

謝辞：本講演の成果の一部は太田和夫先生（電通大）・四方順司先生（横浜国大）との共同研究に基づくものです
本研究は科研費補助金 JP15H02710, JP17H01752 の助成を受けています

はじめに：暗号理論の歴史から

1. 1950 (1945): 暗号に関する最初の理論的成果 (Shannon)
   - ☞ 完全秘匿性（情報理論的安全性）
2. 1976: 公開鍵暗号・電子署名のアイデア (Diffie–Hellman)
   - ☞ 一方向性・計算量的安全性
3. 1978: 初めての公開鍵暗号・電子署名 (Rivest–Shamir–Adleman)
4. 1982: 公開鍵暗号の安全性概念の整備 (Goldwasser–Micali)
   - ☞ 識別不可能性
   - ☞ 強秘匿性

疑問

- 安全性概念 (security notion) とは？

情報理論的安全性概念

（ボンヤリとした）定義

- 攻撃者の計算能力に依存しない安全性

例 (使い捨て暗号 (One-Time Pad, OTP))

完全秘匿性，ε–秘匿性

（ボンヤリとした）定義：秘密鍵暗号の場合

- 平文と暗号文の確率的独立性

1. なぜ確率的独立性なの？
   - ☞ よく言われる，無制限（無限）の計算能力とは？
     - 計算量的（≈ 暗号理論的な）安全性との関係は？
   - ☞ 独立性の尺度：情報量で定義すると情報理論的？
     - 変動距離じゃダメ？
     - 情報が漏れる情報理論的暗号方式（例えばランプ型秘密分散法）に対する攻撃手法は？
2. 安全性の定義と暗号プリミティブの関係は？
   - ☞ いつでも実現できるわけではない（例：情報理論的に安全な PKC）

本講演でお話ししたいこと

1. 情報理論的安全性の意味づけの多角的考察
   - ☞ どのような意味で,
     無制限の計算能力をもつ攻撃者に対して安全といえるか？
   - ☞ 計算量的安全性との関係
2. 安全性と暗号プリミティブの微妙な関係
   - ☞ 推測秘匿性のもとでの秘密鍵暗号・秘密分散法
     - ▶ 実現できる安全性が微妙に異なる

お伝えしたいこと

- ▶ 安全性の定式化は難しく，ときに繊細.

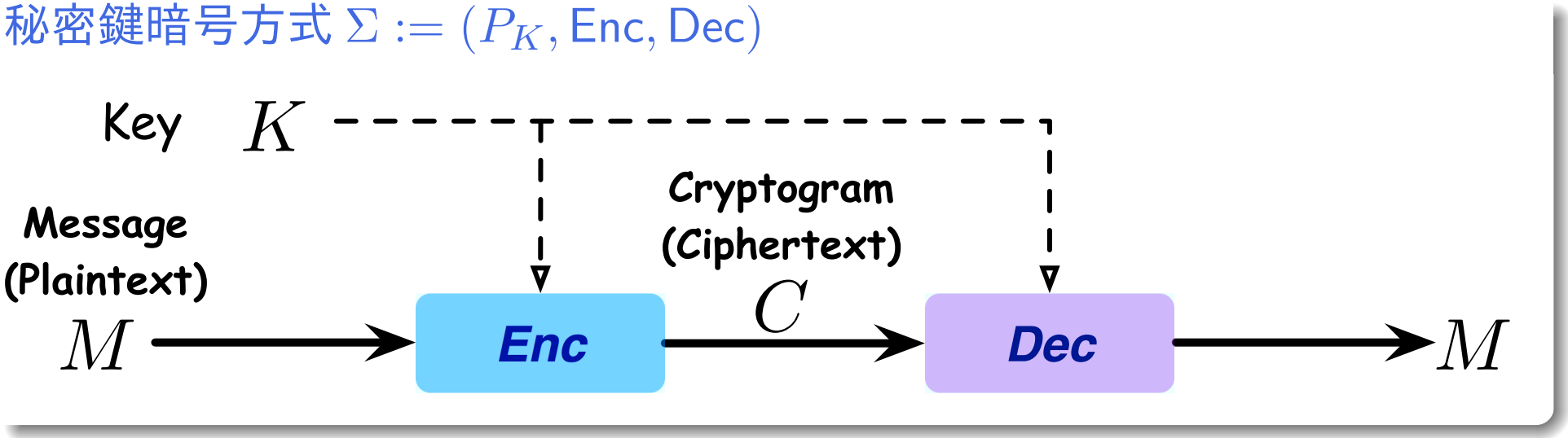
目次

《情報理論的安全性の多角的考察》

完全秘匿性とその解釈

秘密鍵暗号方式

秘密鍵暗号方式 $\Sigma := (P_K, \mathrm{Enc}, \mathrm{Dec})$



- $P_M \in \mathcal{P}(\mathcal{M})$, $P_K \in \mathcal{P}(\mathcal{K})$
- $\mathrm{Enc} : \mathcal{M} \times \mathcal{K} \longrightarrow \mathcal{C}$
- $\mathrm{Dec} : \mathcal{C} \times \mathcal{K} \longrightarrow \mathcal{M}$

完全秘匿性

定義 (Perfect Secrecy, [Shannon])

$\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ が完全秘匿 (perfect secrecy) を達成 $\overset{\text{def}}{\iff}$

$$\forall P_M \in \mathcal{P}(\mathcal{M}), \forall c \in \mathcal{C},\ \forall m \in \mathcal{M},\ P_{M|C}(m|c) = P_M(m)$$

- 平文 M と暗号文 C の確率的独立性
- "a posteriori probability (of M)" $=$ "a priori probability" (of M)

- Shannon は $\forall P_M \in \mathcal{P}(\mathcal{M})$ とは書いていないような…

完全秘匿性 = 「最強」の安全性？

定義 (Perfect Secrecy, [Shannon])

$\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ が完全秘匿 (perfect secrecy) を達成 $\overset{\text{def}}{\iff}$

$$\forall P_M \in \mathcal{P}(\mathcal{M}), \forall c \in \mathcal{C},\ \forall m \in \mathcal{M},\ P_{M|C}(m|c) = P_M(m)$$

- One-time pad (Vernam Cipher) はこの条件を満たす
- 平文 M と暗号文 C の確率的独立性

  $\Longrightarrow$ 無制限の計算能力をもつ攻撃者に対して安全 = “最強の” 安全性

疑問

- なぜ，なにをもって「最強」と言えるのか？
- 攻撃者はどこにいるのか？

一つの解釈：条件付き確率と写像の関係

○条件付き確率は写像（関数）の一般化

$$f : \mathcal{X} \to \mathcal{Y} \text{ に対して } \quad P_{Y|X}(y|x) = \begin{cases} 1 & \text{if } y = f(x) \\ 0 & \text{if } y \neq f(x) \end{cases}$$

- $P_{Y|X}(y|x) = 1$ となる (x, y) を列挙すれば，f で関係づけられた (x, y) の表が出来る

Example (素因数分解)

$f : \mathbb{N} \to \mathbb{N}$ s.t. $f(x) :=$ $x \in \mathbb{N}$ を素因数分解したときの最大素因数

$$Y = f(X) \implies P_{Y|X}(y|38) = \begin{cases} 1 & \text{if } y = 19 \\ 0 & \text{if } y \neq 19 \end{cases}$$

独立性の尺度は？：ε-秘匿性

定義 (独立性の尺度の例)

- 相互情報量：$I(C;M) \leq \varepsilon$
- 変動（統計）距離：$d(P_{CM}, P_C P_M) \leq \varepsilon$

cf)

$$I(X;Y) := H(X) - H(X|Y)$$

$$d(P_X, P_Y) := (1/2) \sum_a |P_X(a) - P_Y(a)|$$

[Remark] Pinsker's の不等式：

$$d(P_{CM}, P_C P_M)^2 \leq \frac{1}{2 \ln 2} I(C;M)$$

疑問

- なぜ変動距離が大事なのか？

攻撃者の存在

暗号理論では…

- 攻撃者の挙動や能力が安全性概念の定式化に不可欠
  - ☞ 強秘匿性 (Semantic security)
  - ☞ 識別不可能性 (Indistinguishability)

疑問

- 完全秘匿性，ε-秘匿性との関係は？

計算量的安全性との関係

統計的 ε–強秘匿性 (Semantic Security)

定義 ([Goldwasser–Micali], [Russell–Wang])

$\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ が統計的 ε–強秘匿性 ($\mathsf{SS}(\varepsilon)$) を満たす $\overset{\text{def}}{\iff}$

$$\underline{\forall P_M} \in \mathcal{P}(\mathcal{M}), \forall f : \mathcal{C} \to \{0,1\}, \forall h : \mathcal{M} \to \{0,1\}, \exists G_f,$$

$$\left|\Pr\{f(C) = h(M)\} - \Pr\{G_f = h(M)\}\right| \leq \varepsilon$$

where G_f: f に依存するが M と独立な 0/1–確率変数

- 直感的意味:
  - ☞ f : 暗号文 C から $h(M)$ (=平文から得られる 1bit 情報) を推測
  - ☞ G_f : f は知っているが暗号文を使わずに $h(M)$ を推測
    $\Longrightarrow$ C を知っていても ($f(C)$), 知らなくても (G_f) 推測精度は同じ (ε)
- 計算量的安全性 : ε, f, g に計算量制約をかける

統計的な ε–識別不可能性

定義

$\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ が完全秘匿 (perfect secrecy) を達成 $\overset{\text{def}}{\iff}$

$$\forall m_0, \forall m_1 \in \mathcal{M}, \quad \forall f : \mathcal{C} \to \{0,1\},$$
$$\left| \Pr\{f(C) = 1 \mid M = m_0\} - \Pr\{f(C) = 1 | M = m_1\} \right| = 0$$

- 直観：
  - ☞ f は「c は m_0 と m_1 のどちらの暗号文か」答える (K を使わない)
  - ☞ c が m_0 と m_1 のどちらを暗号化したか区別できない
- 計算量的安全性との関係：
  - ☞ $\begin{cases} \text{“}= 0\text{”} \Rightarrow \text{“}\leq \varepsilon\text{”} \ (= 1/\mathsf{poly}(n)) \\ f \text{ を多項式サイズ circuit 族に制約} \end{cases} \Longrightarrow$ 計算量的識別不可能性

識別不可能性と強秘匿性

（ザックリとした）定理

- 識別可能性と強秘匿性は等価である

詳しく知りたい方は…

- 計算量的安全性
  - ☞ 公開鍵暗号：Goldwasser–Micali, J. of Computer & Sys. Sci., 1982
    - Goldreich: Foundations of Cryptography vol.2, Cambridge に詳しい
  - ☞ 常に成り立つとは限らないことも知られている
- 情報理論的安全性
  - ☞ 盗聴通信路：Bellare–Tessaro–Vardy, CRYPTO 2012
  - ☞ 共通鍵暗号・鍵共有：Iwamoto–Ohta–Shikata, IEEE-IT, Jan. 2018
    - 初出は ISIT2011 （共通鍵暗号）& 2013（鍵共有）

完全秘匿性のバリエーションとその相互関係

変動距離による識別不可能性の定義

- よく知られた事実：

$$d(P_X, P_Y) = \max_{f:\mathcal{X}\to\{0,1\}} \left|\Pr\{f(X)=1\} - \Pr\{f(Y)=1\}\right|$$

定理 (識別不可能性の等価な定義)

$\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ が ε–識別不可能性 $\mathsf{IND}(\varepsilon)$ を満たす

$$\overset{\text{def}}{\iff} \forall m_0, \forall m_1 \in \mathcal{M}, \quad \forall f : \mathcal{C} \to \{0,1\},$$

$$\left|\Pr\{f(C)=1 \mid M=m_0\} - \Pr\{f(C)=1 \mid M=m_1\}\right| \le \varepsilon$$

$$\iff \forall m_0, \forall m_1 \in \mathcal{M},$$

$$\max_{f:\mathcal{C}\to\{0,1\}} \left|\Pr\{f(C)=1 \mid M=m_0\} - \Pr\{f(C)=1 \mid M=m_1\}\right| \le \varepsilon$$

$$\iff \forall m_0, \forall m_1 \in \mathcal{M}, \quad d(P_{C|M}(\cdot|m_0), P_{C|M}(\cdot|m_1)) \le \varepsilon$$

識別不可能性と完全秘匿性

- IND(0) は次のように書ける：

$$\forall m_0, \forall m_1 \in \mathcal{M}, \quad d(P_{C|M}(\cdot|m_0), P_{C|M}(\cdot|m_1)) = 0$$

- これは完全秘匿性の定義と同値：

$$\forall P_M \in \mathcal{P}(\mathcal{M}), \forall m \in \mathcal{M}, \quad \forall c \in \mathcal{C}, \quad P_{C|M}(c|m) = P_C(c)$$

☞ $\Downarrow$ は簡単，$\Uparrow$ は特殊な平文の分布を仮定

疑問

- 完全秘匿性・ε–秘匿性の定義で，条件に付けるのは m ？ c ？

変動距離に基づく 3 つの情報理論的安全性

定義 (統計距離に基づく 3 つの安全性)

$$\mathrm{PS}_{*M}(\varepsilon)\text{:} \quad \forall P_M \in \mathcal{P}(\mathcal{M}), \forall c \in \mathcal{C}, \qquad d(P_{M|C}(\cdot|c), P_M(\cdot)) \leq \varepsilon$$

$$\mathrm{PS}_{C*}(\varepsilon)\text{:} \quad \forall P_M \in \mathcal{P}(\mathcal{M}), \forall m \in \mathcal{M}, \qquad d(P_{C|M}(\cdot|m), P_C(\cdot)) \leq \varepsilon$$

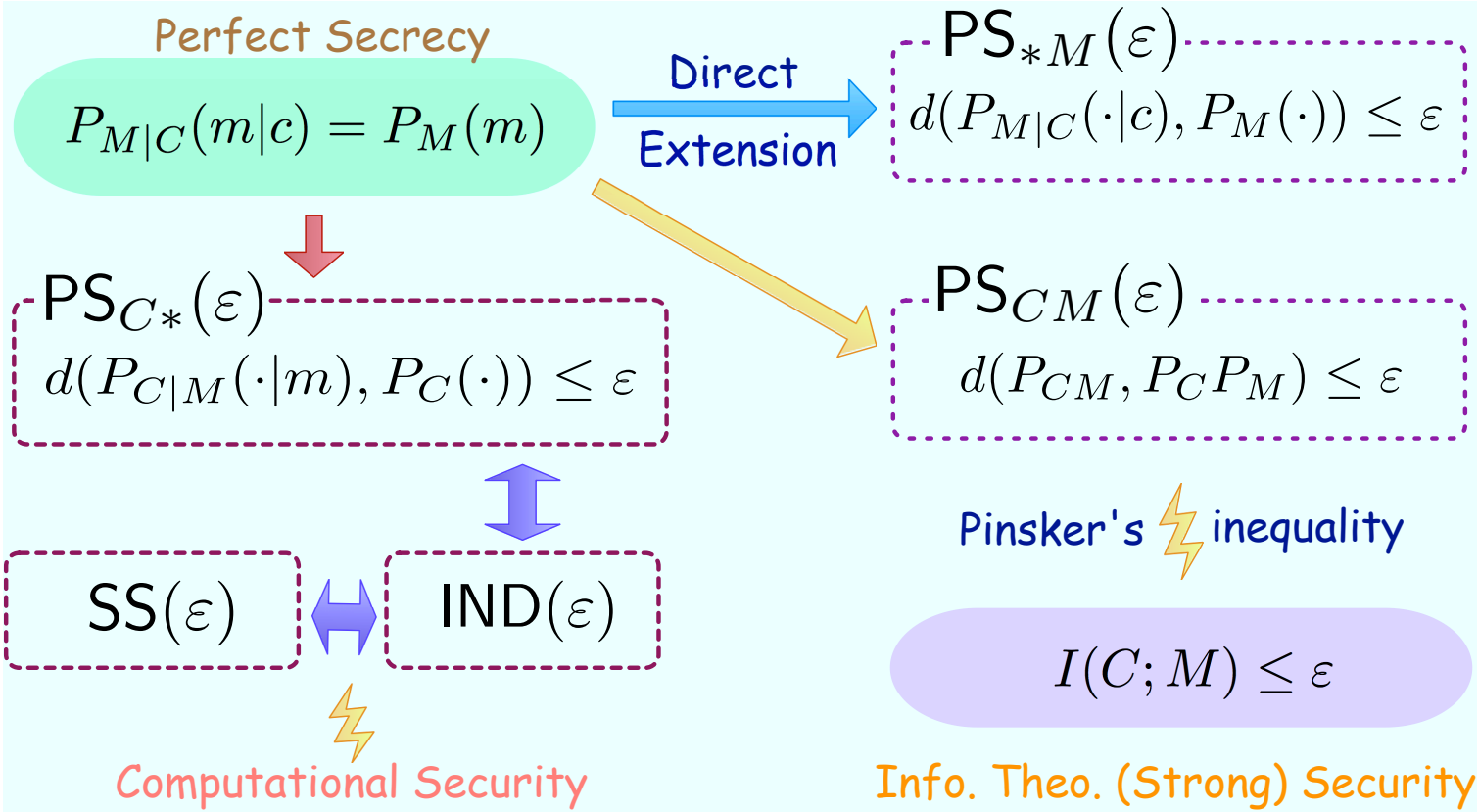
$$\mathrm{PS}_{CM}(\varepsilon)\text{:} \quad \forall P_M \in \mathcal{P}(\mathcal{M}), \qquad d(P_{CM}(\cdot|\cdot), P_C(\cdot)P_M(\cdot)) \leq \varepsilon$$

- $\varepsilon = 0$ の場合, 3 つの安全性基準は等価
  - ☞ $\varepsilon \neq 0$ の場合, 3 つの安全性基準の関係は?

疑問

1. 3 つの安全性の関係
2. (統計的) 識別不可能性 IND(ε)，強秘匿性 SS(ε) との関係

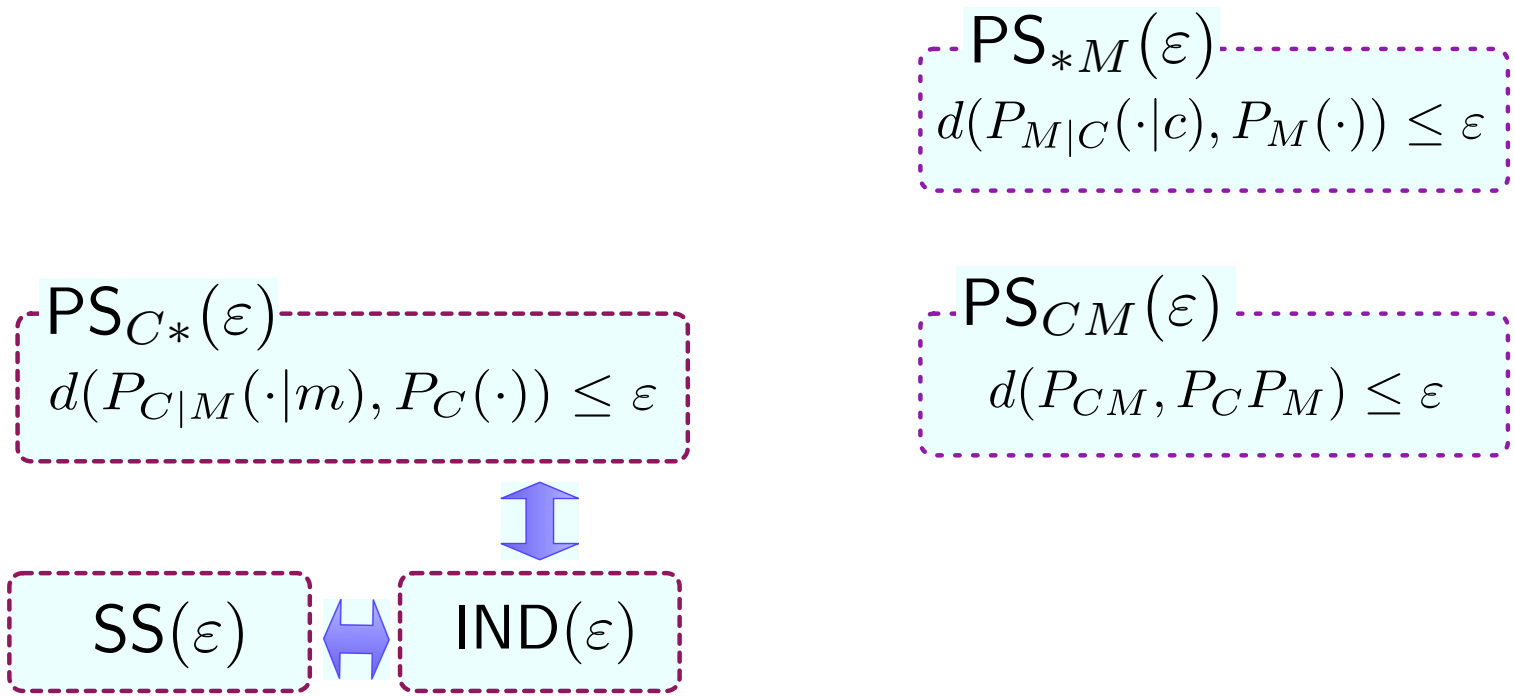
ここまでのまとめ



- $\varepsilon = 0$ なら，全ての安全性概念は等価．$\varepsilon > 0$ の場合は？

幾つかの情報理論的安全性

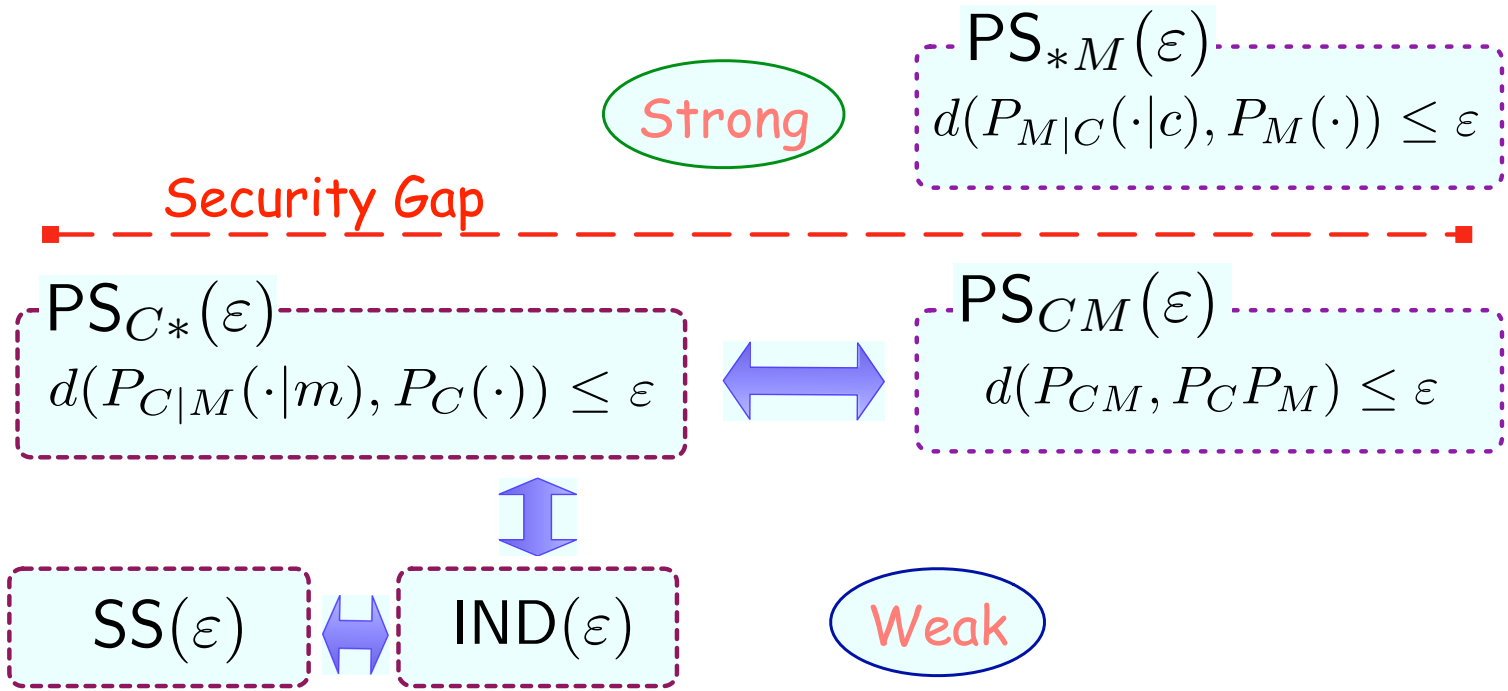
- 少なくとも以下の 5 つの安全性が考えられる



M. Iwamoto, K. Ohta, and J. Shikata,
"Security Formalizations and Their Relationships for Encryption and Key Agreement in Information-Theoretic Cryptography,"
IEEE Trans. Information Theory, vol. 64, issue 1, pp. 654–685, 2018. DOI: `https://doi.org/10.1109/TIT.2017.2744650`

安全性の間の関係

- 実は，任意の c に対して安全性を保証する定義が強い



M. Iwamoto, K. Ohta, and J. Shikata,
"Security Formalizations and Their Relationships for Encryption and Key Agreement in Information-Theoretic Cryptography,"
IEEE Trans. Information Theory, vol. 64, issue 1, pp. 654–685, 2018. DOI: `https://doi.org/10.1109/TIT.2017.2744650`

安全性ギャップの考察

秘密鍵暗号方式 $\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ と確率遷移行列

定理 ((K, Enc) と $\{P_{C|M}(c|m)\}_{c,m}$)

- M, K は独立 $\Rightarrow$ Enc と P_K から $P_{C|M}(\cdot|\cdot)$ が定まる

証明：

$$
\begin{aligned}
P_{CM}(c,m) =& \Pr\{C = c, M = m\} \\
=& \Pr\{\mathsf{Enc}(M,K) = c, M = m\} \\
=& \sum_{k:\mathsf{Enc}(m,k)=c} P_{MK}(m,k) \\
=& P_M(m) \sum_{k:\mathsf{Enc}(m,k)=c} P_K(k) \qquad \because) M \perp K \\
=& P_M(m) \Pr\{\mathsf{Enc}(m,K) = c\},
\end{aligned}
$$

- 逆は？

例: 鍵が非一様な場合の One time pad

例 (鍵が非一様分布な OTP)

$P_K(0) = 0.2$ であるような one-time pad:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ (i.e., } \oplus 0) \text{ with prob. } 0.2, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \text{ with prob. } 0.8$$

- 確率遷移行列 $\mathbb{P}_{C|M}$: 右から確率分布 P_M を入力すると思ってください

$$0.2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 0.8 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix}$$

- 入力が $\mathbb{P}_M = {}^t[0.3, 0.7]$ なら，出力分布 $\mathbb{P}_C$ は

$$\mathbb{P}_C = \begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix} \begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix} = \begin{bmatrix} 0.62 \\ 0.38 \end{bmatrix}$$

秘密鍵暗号方式 $\Sigma = (P_K, \mathsf{Enc}, \mathsf{Dec})$ と確率遷移行列

定理 ((K, Enc) と $\{P_{C|M}(c|m)\}_{c,m}$)

- $|\mathcal{C}| = |\mathcal{M}|$ とする.
- $k^* \in \mathcal{K}$ を固定すると，$\mathsf{Enc}(\cdot, k^*) : \mathcal{M} \to \mathcal{C}$ は全単射

$$\left\{P_{C|M}(c|m)\right\}_{\substack{c\in\mathcal{C}\\ m\in\mathcal{M}}} \longleftrightarrow \sum_{k\in\mathcal{K}} P_K(k)\cdot\Pi_k$$

where Π_k : permutation matrix

これは二重確率行列. 逆もいえる (Birkoff–von Neuman Theorem)

定理 (Birkoff–von Neuman Theorem)

- 行列 A が二重確率的 $\Leftrightarrow$ 行列 A が置換行列の凸結合

行列 $\mathbb{P}_{C|M}$ と識別不可能性

定義 (IND(ε))

$$\forall m_0, \forall m_1 \in \mathcal{M}, \quad \forall c \in \mathcal{C}, \quad d(P_{C|M}(\cdot|m_0), P_{C|M}(\cdot|m_1)) \leq \varepsilon$$

例 (One time pad)

- $\mathbb{P}_{C|M}$ の列同士の変動距離の最大値を ε にする
- $n = |\mathcal{X}|$

$$\mathbb{P}_{C|M} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{array} \begin{array}{c} \begin{array}{cccccc} m_1 & m_2 & \cdots & m_{n-1} & m_n \end{array} \\ \begin{bmatrix} n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \\ n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \\ n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \end{bmatrix} \end{array} \Longrightarrow \varepsilon = 0$$

$\mathsf{PS}_{*M}(\varepsilon)$ と他の安全性とのギャップ

例：$\delta = \varepsilon/2 \in (0, n^{-1}]$, n: even

$$\mathbb{P}_{C|M} = \begin{array}{c} \\ c_1 \\ c_2 \\ c_{n-1} \\ \vdots \\ c_n \end{array} \begin{array}{c} \begin{array}{cccc} m_1 & m_2 & \cdots & m_{n-1} \end{array} \quad m_n \\ \begin{bmatrix} n^{-1}+\delta & n^{-1}-\delta & \cdots & n^{-1}+\delta & n^{-1}-\delta \\ n^{-1}-\delta & n^{-1}+\delta & \cdots & n^{-1}-\delta & n^{-1}+\delta \\ n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n^{-1} & n^{-1} & \cdots & n^{-1} & n^{-1} \end{bmatrix} \end{array}$$

[Remark]

- $\mathbb{P}_{C|M}$ は二重対角行列
- $\mathsf{IND}(2\delta)$ を満たす．つまり $\mathsf{SS}(2\delta)$, $\mathsf{PS}_{C*}(2\delta)$, $\mathsf{PS}_{CM}(2\delta)$ も満たす．
- $\delta > 0$ は任意に小さく取れる (e.g., $\delta = 1/n$)

$\mathsf{PS}_{*M}(\varepsilon)$ が一番強い

- P_M を一様分布として，$\delta = n^{-1}$ ととると，

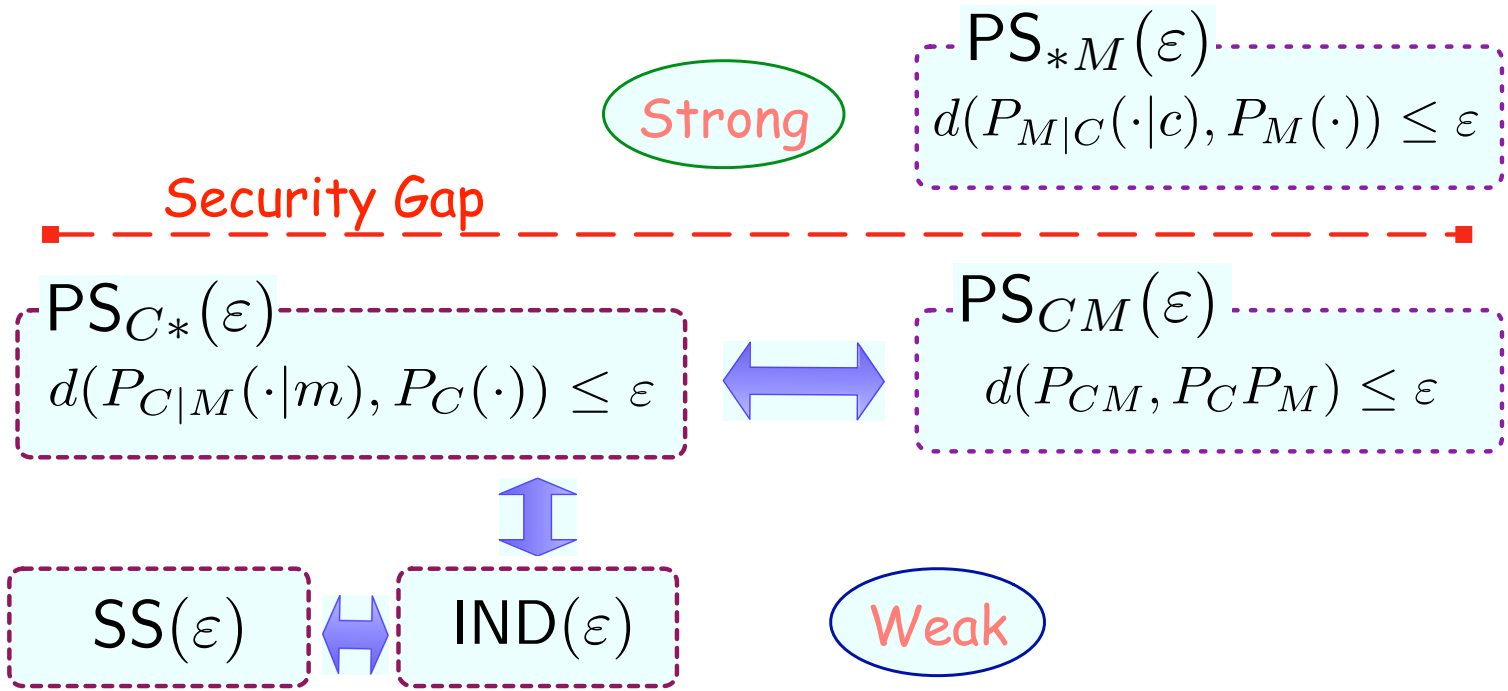
$$\mathbb{P}_{M|C} = \mathbb{P}_{C|M}^{\top} = \begin{array}{c} \\ m_1 \\ m_2 \\ m_3 \\ \vdots \\ m_n \end{array} \begin{array}{c} \begin{array}{ccccc} c_1 & c_2 & \cdots & c_{n-1} & c_n \end{array} \\ \begin{bmatrix} 2n^{-1} & 0 & \cdots & n^{-1} & n^{-1} \\ 0 & 2n^{-1} & \cdots & n^{-1} & n^{-1} \\ 2n^{-1} & 0 & \cdots & n^{-1} & n^{-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 2n^{-1} & \cdots & n^{-1} & n^{-1} \end{bmatrix} \end{array}$$

$$\Longrightarrow d(P_{M|C}(\cdot|c), P_M(\cdot)) = \begin{cases} 1/2, & \text{if } c = c_1 \text{ or } c_2 \quad [\text{with prob. } 1/n] \\ 0, & \text{otherwise} \end{cases}$$

- これは $\mathsf{PS}_{*M}(1/2)$ 安全
- $\mathsf{PS}_{*M}(\varepsilon)$ は安全でない暗号文 c を（確率が小さくても）許さない

まとめ：安全性の間の関係（再掲）

- 実は，任意の c に対して安全性を保証する定義が強い



M. Iwamoto, K. Ohta, and J. Shikata,
"Security Formalizations and Their Relationships for Encryption and Key Agreement in Information-Theoretic Cryptography,"
IEEE Trans. Information Theory, vol. 64, issue 1, pp. 654–685, 2018. DOI: `https://doi.org/10.1109/TIT.2017.2744650`

《安全性と暗号プリミティブの微妙な関係》

推測秘匿性のもとでの共通鍵暗号・秘密分散法

推測秘匿性

共通鍵暗号：m: 平文, c: 暗号文

完全秘匿性: 統計的独立性による定義 [Shannon, 1945/1949]

$$\forall P_M \in \mathcal{P}(\mathcal{M}), \forall m, \forall c, P_{M|C}(m|c) = P_M(m)$$

推測秘匿性： 推測成功確率に基づく

- Average Guessing Secrecy, A-GS: [Alimomeni, Safavi-Naini, ICITS2012]

$$\mathbb{E}_C \left[\max_m P_{M|C}(m|C) \right] = \max_m P_M(m)$$

- Worst-case Guessing Secrecy, W-GS: [I–Shikata, ICITS2013]

$$\max_{m,c} P_{M|C}(m|c) = \max_m P_M(m)$$

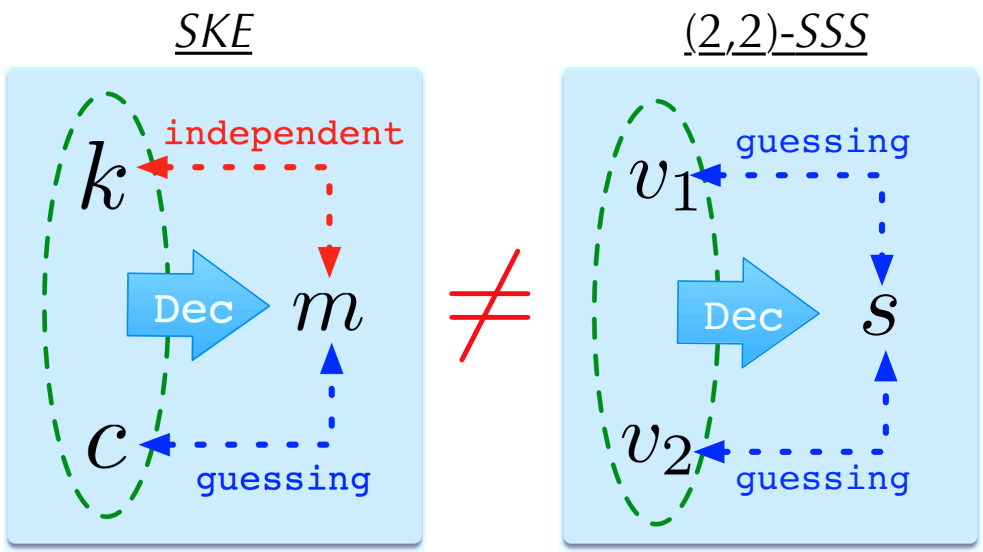
- 明らかに [weaker] A-GS $\preceq$ W-GS $\preceq$ PS [stronger]

安全性概念の関係がプリミティブによって違う

- 安全性概念の関係はプリミティブによって異なる [‡]:
  - ☞ 共通鍵暗号： A-GS $\prec$ W-GS $=$ PS
  - ☞ 秘密分散法： A-GS $\prec$ W-GS $\prec$ PS

  (“$\prec$” は非自明なギャップの存在を意味する)

- 理由：



[‡] M. Iwamoto and J. Shikata, ICITS2013, ISIT2014, ISIT2015, with recent result.

本講演のまとめ

1. 情報理論的安全性の意味づけの多角的考察
   - ☞ どのような意味で，
     無制限の計算能力をもつ攻撃者に対して安全といえるか？
   - ☞ 計算量的安全性との関係
2. 安全性とプリミティブが微妙な関係にある例
   - ☞ 推測秘匿性のもとでの秘密鍵暗号・秘密分散法

お伝えしたかったこと

- 安全性の定式化は難しく，ときに繊細.

Fin.

ありがとうございました