教養原論 数理の世界
(数学とコンピュータ)
RSA 暗号

野呂 正行

神戸大学大学院理学研究科数学専攻

November 7, 2013

暗号とは

- **暗号化**
  ある規則で, 文字を別の文字に変換
- **復号**
  暗号文をもとの文 (平文; ひらぶん) に戻す
- **単純なもの : 表を使う**
  例 : アルファベットを単に何文字かずらす (シーザー暗号)
  ⇒ 文字の出現頻度を調べると解読できる
  参考 : 英文字の出現頻度 (COD; **数字は%**)
  E(11), A(8.5), R(7.6), I(7.5), O(7.2), T(7.0), N(6.7), S(5.7), L(5.5), C(4.5), U(3.6), D(3.4), P(3.2), M(3.0), H(3.0), G(2.5), B(2.1), F(1.8), Y(1.8), W(1.3), K(1.1), V(1.0), X(0.3), Z(0.3), J(0.2), Q(0.2)

暗号方式

暗号は, 公開されたアルゴリズムにより暗号化, 復号

⇒ 秘密保持のため, 鍵が必要

- **鍵とは**

  暗号化, 復号アルゴリズムの実行に必要なデータ

- **共通鍵暗号** (例 : DES, AES)

  暗号化, 復号に同じ鍵を使う

  鍵は, 通信者 A, B が何らかの方法で共有する

  ⇒ **鍵の共有方法が問題**

- **公開鍵暗号** (例 : RSA, 楕円曲線暗号)

  1. A は B が公開している鍵でデータを暗号化して送る
  2. B は, 自分が持つ秘密鍵で復号する

  ⇒ **一時的に使う秘密鍵の送付に使える**

例 : https://. . .

A が web shop B で物を買いたいとする
(B の URL は信用できるとする)
問題点 : インターネット上では常に盗聴の危険あり

- 初対面の相手と安全に通信するにはどうするか?
  データは暗号化する必要がある
  ⇒ 自分と相手だけが知る鍵が必要
- 初対面で, どうやって 鍵を共有するか?
  A が共通鍵を作り, B に渡せばよい
- 共通鍵を相手に安全に渡すにはどうすればよいか
  B が公開鍵をつくり, 公開し, A は共通鍵を, その公開鍵で暗号化して渡す

公開鍵暗号の要件

- **公開鍵から秘密鍵がバレてはいけない**
  暗号化方法は当然公開されている (復号方法も)
  公開鍵と秘密鍵は当然密接に関係している
  $\Rightarrow$ **公開鍵から秘密鍵がバレないような方式が必要**

- **暗号文から平文が簡単に復元できない**
  鍵がなくても平文が復元できるようでは意味がない
  例 : 一文字ずつ変換したのでは, 頻度解析などで解読できる
  $\Rightarrow$ **複数文字 (文字列) を変換できなければいけない**

RSA 暗号

Rivest, Shamir, Adleman により考案された (1977)
前述の要件を満たす暗号方式の一つ

- **公開鍵 n は素数 p, q を 2 つかけたもの** $(n = pq)$
  例 : 1024 ビット RSA 暗号の公開鍵は, 150 桁程度の素数を 2 つかけたもの
  $0 \leq x \leq n - 1$ なる整数 x を, 同じ範囲の整数に変換, 逆変換できる.

- **文字列の変換**
  各文字を数に変換してつなげた大きい整数として変換
  アルファベット (ASCII) : 1 バイト
  漢字, かな : JIS なら 2 バイト, UTF-8 なら 3 バイト
  以上

参考：文字コードについて

- ASCII
  アルファベット, 数字, 記号を 1 バイト (0-127) で表す.
- JIS コード (ISO-2022-JP)
  漢字を ASCII 2 バイトで表すが, 通常の ASCII と混在するために, 表の切り替え (エスケープシークエンス) が必要.
- EUC-JP
  JIS 2 バイトの各バイトの最上位ビットを 1 にしたもの. ASCII コードは最上位ビットが 0 なので, 切り替えなしに混在可能.
- UTF-8 (Unicode)
  世界中のすべての文字を表現する表. ASCII は 1 バイト, その他の文字を 2 から 6 バイトで表す. 漢字は 3 バイト以上必要.

RSA 暗号：公開鍵 (n, e), 秘密鍵 d

- **公開鍵** (n, e) ($e = 65537$ **がよく用いられる**)
  大きい素数 p, q ($p \neq q$) **を選ぶ**
  実用上 p, q **は** 10^{150} **くらいで**, $p-1, q-1$ **が** $e = 65537$ **で割り切れないものを選ぶ**.
  作り方 : **乱数を発生させ**, **素数テスト**
  ($p-1$, $p+1$ **が小さい素因数で分解されない**, **などいろいろ要件がある**)
  $n = pq$ **に対し**, (n, e) **が公開鍵**
- **秘密鍵** d
  e **は** $n' = (p-1)(q-1)$ **を割り切らない素数なので**,
  $GCD(e, n') = 1$
  $\Rightarrow de \equiv 1 \bmod n'$ **を満たす** $d > 0$ **がある**
  この $d > 0$ **を秘密鍵とする**.
  d **の計算法** : **拡張ユークリッド互除法**

RSA 暗号：安全性

- **鍵の安全性**

  d を作るには $n' = (p-1)(q-1)$ が必要

  $\Rightarrow$ p, q が必要 (たぶん)

  $\Rightarrow$ n の素因数分解が必要

  $\Rightarrow n \simeq 10^{300}$ では事実上無理

  (最近では $\Rightarrow n \simeq 10^{600}$ が推奨されている)

  すなわち，「数学者，計算機科学者，コンピュータの現在の能力では素因数分解ができない」ことに基づく暗号方式が RSA 暗号

  (量子コンピュータが実用化されると危うい?)

- **e の選び方**

  e はある程度大きい素数ならなんでもよい

  根拠：$a^e \bmod n$ から a を復元することは困難 (後述)

RSA 暗号：暗号化 — $E = A^e \bmod n$

1. **データを数字の列に変換**

   例えば, UTF-8 3 **バイト文字列なら**

a_1	a_2	$\cdots$	a_l

   $0 \le a_i \le 256^3 - 1$ (3 **バイト**)

2. 1. **で作った数の列を**, n **未満の数に分割する**.

a_1	$\cdots$	a_m	a_{m+1}	$\cdots$	a_{2m}	$\cdots$

   3 **バイトが** m **文字分** $\Rightarrow 256^3$ **進** m **桁**

   $\Rightarrow 256^{3m} \le n$ **となるように** m **をとればよい**.

   $A_1 =$ | a_1 | $\cdots$ | a_m |, $A_2 =$ | a_{m+1} | $\cdots$ | a_{2m} |, $\cdots$ **とおく**.

   すなわち,

   $$A_1 = a_1 \cdot 256^{3(m-1)} + \cdots + a_{m-1} \cdot 256^3 + a_m, \ldots$$

   と見なす. **このとき** $0 \le A_i \le n - 1$ **である**.

3. **各** A_i **を暗号化** : $E_i \leftarrow A_i^e \bmod n$

RSA 暗号：復号 — $D = E^d \bmod n$

1. **各 E_i に対し $D_i \leftarrow E_i^d \bmod n$**
   実は, $D_i = A_i$ となっている.

2. **数字の列を元に戻す**
   作った D_i を, 3 バイトずつに区切り, 文字に戻す

注意：E_i から A_i を求めるのは困難.

$a^e \bmod n = E_i$ となる数 a は何か? ($\bmod n$ **での** e **乗根**)

$\Rightarrow 0^e \bmod n, 1^e \bmod n, \ldots, (n-1)^e \bmod n$ **を計算すればいずれ求まるが,** n **が巨大だと無理**

簡単な例

$n = 1223460961843067$, $e = 65537$, $d = 115276424115473$

暗号文：139187734536180 **を復号してみる**.

$$139187734536180^d \equiv 253530592481414 \bmod n$$

$$253530592481414 = (\text{e695b0e79086})_{16}$$

実は $(\text{e695b0})_{16}$ **は「数」**, $(\text{e79086})_{16}$ **は「理」** (UTF-8 **で**)

実際, **「数理」を暗号化してみると**

$$253530592481414^e \equiv 139187734536180 \bmod n$$

と, **上で与えた暗号文になることがわかる**.

数学的な証明が必要な部分

$E = (A^e \bmod n)$, $D = (E^d \bmod n)$, $0 \leq A, D \leq n-1$ **ならば**

$D = A$

これは $A^{ed} \equiv A \bmod n$ **を示せばよい**

基本的事実

p, q **を相異なる素数とする.**

- ab **が** p **で割り切れるなら** a **か** b **が** p **で割り切れる.**
  これが p **が「素」であることの基本的性質**
- a **が** pq **で割り切れる** $\Leftrightarrow$ a **が** p **でも** q **でも割り切れる**
- p **が素数なら** $a^p \equiv a \bmod p$ **(フェルマーの小定理)**
  実際には,「a **が** p **で割り切れないとき** $a^{p-1} \equiv 1 \bmod p$
  」を使う.

$A^{de} \equiv A \bmod n$

$de \equiv 1 \bmod n'$ **より** $de - kn' = 1$ **を満たす** k **がある.**

$A^{de} = A^{1+kn'}$ **より** $A^{1+kn'} \equiv A \bmod n$ **をいえばよい.**

$n = pq$ **より次をいえばよい:**

$$A^{1+kn'} \equiv A \bmod p, \quad A^{1+kn'} \equiv A \bmod q$$

1. $A \equiv 0 \bmod p$ **のとき**

   $A^{1+kn'} \equiv 0 \bmod p$ **より** $A^{1+kn'} \equiv A \bmod p$

2. $A \equiv 0 \bmod p$ **でないとき (**A **が** p **で割り切れないとき)**

   $A^{1+kn'} = A \cdot A^{k(p-1)(q-1)} = A \cdot (A^{p-1})^{k(q-1)}$ **で,**

   A **が** p **で割り切れないから**

   $A^{p-1} \equiv 1 \bmod p$ **(フェルマーの小定理).**

   よって, $A^{1+kn'} \equiv A \bmod p$.

(mod$q$ **も同様)**

$A^e \bmod n$, $E^d \bmod n$ の計算法

繰り返し 2 乗法 (repeated squaring)

剰余計算の基本 : $(ab) \bmod n = (a \bmod n)(b \bmod n) \bmod n$

と指数法則 : $a^{u+v} = a^u \cdot a^v$

を繰り返し使う.

1. $k = (k_m k_{m-1} \cdots k_0)_2$ **と 2 進で書く**
2. $a_i = a^{2^i} \bmod n \ (i = 0, 1, \ldots, m)$ **を計算する**
   $a^{2^{i+1}} = a^{2^i + 2^i} = a^{2^i} \cdot a^{2^i}$ **より** $a_{i+1} = a_i^2 \bmod n$
3. $k_i = 1$ **である** i **に対する** a_i **を全部かけて** mod$n$ **する**
   例: $k = 53 = 2^5 + 2^4 + 2^2 + 2^0 = (110101)_2$ **なら**,
   $a^k = a^{2^5} \cdot a^{2^4} \cdot a^{2^2} \cdot a^{2^0}$ **より**
   $a^k \bmod n = (a_5 \cdot a_4 \cdot a_2 \cdot a_0) \bmod n$

例: $12^{100} \bmod 15$ の計算

$100 = 2^6 + 2^5 + 2^2 = (1100100)_2$

$a_0 = 12 \bmod 15 = 12$

$a_1 = 12^2 \bmod 15 = 9$

$a_2 = 9^2 \bmod 15 = 6$

$a_3 = 6^2 \bmod 15 = 6$

$a_4 = 6^2 \bmod 15 = 6$

$a_5 = 6^2 \bmod 15 = 6$

$a_6 = 6^2 \bmod 15 = 6$

$12^{100} \bmod 15 = (6 \cdot 6 \cdot 6) \bmod 15$

$= ((6 \cdot 6) \bmod 15 \cdot 6) \bmod 15 = (6 \cdot 6) \bmod 15 = 6$