187 - Day 06

Friday, April 13, 2018 7:50 AM

Day 06 – Merkle-Hellman Knapsack cipher

Hill cipher

Definition 1. *Hill cipher is a* **block cipher** *where* **pairs of plaintext letters** *are encrypted by the transformation*

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mod 26$$

A: 2x2 matrix
that is invertible mod 26.

where A is a 2×2 invertible matrix modulo 26.

Decryption is given by $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \cancel{X}\, A^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \mod 26$

A^{-1} is the inverse of A
in mod 26.

Remark:

. Extend the algorithm to any modulo N
where N = # letters in the alphabet.

. Use any n x n invertible matrix A.
(under mod N)

↳ To find A^{-1} mod N: use a modified Gauss-Jordan elim.

$$[A \mid I] \xrightarrow{\text{GJ elim.}} [I \mid A^{-1}] \mod N.$$

(see my MatLab m-file)

1

Euclidean algorithm - Berlekamp version

The Euclidean algorithm is the process which yields the greatest common divisor d of two given integers A and B where $A > B$. The Berlekamp's algorithm is a slight modification of the Euclidean algorithm that can also give us the inverse of B mod A. → B^{-1} under mod A

Algorithm:

1. Set $r_{-2} = A, r_{-1} = B, p_{-2} = 0, p_{-1} = 1, q_{-2} = 1, q_{-1} = 0$
2. For each $k = 0, 1, 2, \dots$ compute the following

$$r_{k-2} = a_k r_{k-1} + r_k$$
$$p_k = a_k p_{k-1} + p_{k-2}$$
$$q_k = a_k q_{k-1} + q_{k-2}$$

3. Stop at step n where $r_n = 0$

If $\gcd(A,B) \neq 1$ then B^{-1} mod A D.N.E.

Then $r_{n-1} = \gcd(A, B)$. Furthermore, in the case $\gcd(A, B) = 1$ then

$$B \cdot \underbrace{(-1)^n p_{n-1}}_{B^{-1} \text{ under mod } A.} = 1 + A \cdot (-1)^n q_{n-1}$$

Example. Find the inverse of 115 in modulo 12659.

k	r	a	p	q
-2	12659 (A)		0	1
-1	115 (B)		1	0
0	9	110	110	1
1	7	12	1321	12
2	2	1	1431	13
3	1	3	5614	51
4	0	2	12,659	115

gcd(A,B)

Stop →

$\boxed{n=4}$

A B

Here,

$$B^{-1} = (-1)^n p_{n-1}$$
$$= (-1)^4 \cdot 5614$$
$$= 5614$$

In Wolfram Alpha :

>> Quotient Remainder [A , B]

output (a , r)

quotient remainder

Steps of Berlekamp: $\underline{k=0}$

- $r_{-2} = \underline{a_0} \cdot r_{-1} + \underline{r_0}$

$12,659 = \underbrace{110}_{a_0} \times 115 + \underbrace{9}_{r_0}$

- $p_0 = 110 \times 1 + 0 = 110$

$q_0 = 110 \times 0 + 1 = 1$

$\underline{k=1}$: $\underset{r_{-1}}{115} = \underset{a_1}{(12)} \times \underset{r_0}{9} + \underset{r_1}{(7)}$ $\quad a_1 = 12,\ r_1 = 7$

$p_1 = \underbrace{12}_{a_1} \times \underbrace{110}_{p_0} + \underbrace{1}_{p_{-1}} = 1321$

$q_1 = 12 \times \underbrace{1}_{q_0} + \underbrace{0}_{q_{-1}} = 12$

$\underline{k=2}$: $r_0 = a_2 \times r_1 + r_2$ $\quad a_2 = 1,\ r_2 = 2$

$9 = 1 \times 7 + 2$

$p_2 = a_2 \times p_1 + p_0 = 1 \times 1321 + 110 = 1431$

$q_2 = a_2 \times q_1 + q_0 = 1 \times 1 + 12 = 13$

Subset-Sum Problem (SSP)

Given an increasing sequence $a_1 < a_2 < \ldots < a_n$ and a "target" number M. We want to determine if there is a subsequence of the a_i's whose sum is t.

NP-hard

Formally, we want to find the values $x_1, x_2, \ldots, x_n$, where each x_i is either 0 or 1 such that

$$x_1a_1 + x_2a_2 + \cdots x_na_n = M.$$

Example. Solve the SSP for $a_1 = 3, a_2 = 5, a_3 = 11, a_4 = 23, a_5 = 51$, and $M = 67$.

$$67 = 51 + 16$$
$$= 51 + 11 + 5$$

$$67 = a_2 + a_3 + a_5 = 0 \cdot a_1 + 1 \cdot a_2 + 1 \cdot a_3 + 0 \cdot a_4 + 1 \cdot a_5.$$

$$x = (0, 1, 1, 0, 1)$$

If the a_i's form a super-increasing sequence then for every right-hand side M, the SSP either has **no** solution or a **unique** solution.

Definition 2. *A* super-increasing sequence *is a sequence where each number in the sequence is **greater than the sum of those preceding it**.*

Formally, $a_1, a_2, \ldots, a_n$ is a super-increasing sequence if for all $1 \leq k \leq n$,

$$a_k > \sum_{i=0}^{k-1} a_i.$$

Example.

1. $(3, 5, 11, 23, 51)$ ✓ $3 > 0$, $5 > 3$, $11 > 5+3$, $23 > 11+5+3$, $51 > 23+11+5+3$
2. ~~$(1, 4, 7, 12, 19)$~~ $12 = 7+4+1$
3. $(13, 18, 35, 72, 155, 301, 595)$ ✓ (You check)

Example. Solve the SSP for $M = 1003$ and the super-increasing sequence

$$(13, 18, 35, 72, 155, 301, 595).$$

$$\begin{aligned}
1003 &= 595 + 408 \\
&= 595 + 301 + 107 \\
&= 595 + 301 + 72 + 35 \\
&= 595 + 301 + 72 + 35 + 0
\end{aligned}$$

$$\begin{aligned}
1003 = {} & 0 \cdot 13 + 0 \cdot 18 + 1 \cdot 35 + 1 \cdot 72 \\
& + 0 \cdot 155 + 1 \cdot 301 + 1 \cdot 595
\end{aligned}$$

$$x = (0, 0, 1, 1, 0, 1, 1)$$

<u>Remark:</u> If the given sequence is not super-increasing, then the solution may not be unique.

Example. Consider $M = 20$ and the sequence $(1, 2, 3, 4, 5, 6, 7)$.

5

The ASCII (American Standard Code for Information Interchange) Table:

Dec	Bin	Hex	Char	Dec	Bin	Hex	Char	Dec	Bin	Hex	Char	Dec	Bin	Hex	Char
0	0000 0000	00	[NUL]	32	0010 0000	20	space	64	0100 0000	40	@	96	0110 0000	60	`
1	0000 0001	01	[SOH]	33	0010 0001	21	!	65	0100 0001	41	A	97	0110 0001	61	a
2	0000 0010	02	[STX]	34	0010 0010	22	"	66	0100 0010	42	B	98	0110 0010	62	b
3	0000 0011	03	[ETX]	35	0010 0011	23	#	67	0100 0011	43	C	99	0110 0011	63	c
4	0000 0100	04	[EOT]	36	0010 0100	24	$	68	0100 0100	44	D	100	0110 0100	64	d
5	0000 0101	05	[ENQ]	37	0010 0101	25	%	69	0100 0101	45	E	101	0110 0101	65	e
6	0000 0110	06	[ACK]	38	0010 0110	26	&	70	0100 0110	46	F	102	0110 0110	66	f
7	0000 0111	07	[BEL]	39	0010 0111	27	'	71	0100 0111	47	G	103	0110 0111	67	g
8	0000 1000	08	[BS]	40	0010 1000	28	(	72	0100 1000	48	H	104	0110 1000	68	h
9	0000 1001	09	[TAB]	41	0010 1001	29	)	73	0100 1001	49	I	105	0110 1001	69	i
10	0000 1010	0A	[LF]	42	0010 1010	2A	*	74	0100 1010	4A	J	106	0110 1010	6A	j
11	0000 1011	0B	[VT]	43	0010 1011	2B	+	75	0100 1011	4B	K	107	0110 1011	6B	k
12	0000 1100	0C	[FF]	44	0010 1100	2C	,	76	0100 1100	4C	L	108	0110 1100	6C	l
13	0000 1101	0D	[CR]	45	0010 1101	2D	-	77	0100 1101	4D	M	109	0110 1101	6D	m
14	0000 1110	0E	[SO]	46	0010 1110	2E	.	78	0100 1110	4E	N	110	0110 1110	6E	n
15	0000 1111	0F	[SI]	47	0010 1111	2F	/	79	0100 1111	4F	O	111	0110 1111	6F	o
16	0001 0000	10	[DLE]	48	0011 0000	30	0	80	0101 0000	50	P	112	0111 0000	70	p
17	0001 0001	11	[DC1]	49	0011 0001	31	1	81	0101 0001	51	Q	113	0111 0001	71	q
18	0001 0010	12	[DC2]	50	0011 0010	32	2	82	0101 0010	52	R	114	0111 0010	72	r
19	0001 0011	13	[DC3]	51	0011 0011	33	3	83	0101 0011	53	S	115	0111 0011	73	s
20	0001 0100	14	[DC4]	52	0011 0100	34	4	84	0101 0100	54	T	116	0111 0100	74	t
21	0001 0101	15	[NAK]	53	0011 0101	35	5	85	0101 0101	55	U	117	0111 0101	75	u
22	0001 0110	16	[SYN]	54	0011 0110	36	6	86	0101 0110	56	V	118	0111 0110	76	v
23	0001 0111	17	[ETB]	55	0011 0111	37	7	87	0101 0111	57	W	119	0111 0111	77	w
24	0001 1000	18	[CAN]	56	0011 1000	38	8	88	0101 1000	58	X	120	0111 1000	78	x
25	0001 1001	19	[EM]	57	0011 1001	39	9	89	0101 1001	59	Y	121	0111 1001	79	y
26	0001 1010	1A	[SUB]	58	0011 1010	3A	:	90	0101 1010	5A	Z	122	0111 1010	7A	z
27	0001 1011	1B	[ESC]	59	0011 1011	3B	;	91	0101 1011	5B	[	123	0111 1011	7B	{
28	0001 1100	1C	[FS]	60	0011 1100	3C	<	92	0101 1100	5C	\	124	0111 1100	7C	|
29	0001 1101	1D	[GS]	61	0011 1101	3D	=	93	0101 1101	5D	]	125	0111 1101	7D	}
30	0001 1110	1E	[RS]	62	0011 1110	3E	>	94	0101 1110	5E	^	126	0111 1110	7E	~
31	0001 1111	1F	[US]	63	0011 1111	3F	?	95	0101 1111	5F	_	127	0111 1111	7F	[DEL]

Merkle-Hellman Knapsack

Alice (**receiver**) and Bob (**sender**) want to communicate using Knapsack.

1. Alice (**receiver**) chooses
   - A **super-increasing sequence** $(a_1, a_2, \ldots, a_n)$,
   - A **prime** $p > \sum_{i=1}^{n} a_i$, and
   - An **encrypting factor** A with $2 \leq A \leq p-1$.

   She will keep these values secret.

2. For each $1 \leq i \leq n$, Alice then compute the sequence b_i's by

   $$b_i = A \cdot a_i \mod p.$$

   Then she sends the sequence b_i's.

3. Now suppose Bob (**sender**) wants to send a binary message $x_1 x_2 \cdots x_n$ of n-bits. He then uses Alice published sequence b_i's and compute

   $$C = x_1 b_1 + x_2 b_2 + \cdots x_n b_n$$

   then he **sends** the value C to Alice.

4. To decrypt Bob's message, Alice computes $M = A^{-1} \cdot C \mod p$ then solves the SSP

   $$x_1 a_1 + x_2 a_2 + \cdots x_n a_n = M$$

   to recover the binary string $x_1 x_2 \cdots x_n$

Alice	Eve → Public	Bob
$(a_1, \ldots, a_n)$ super-increasing		
prime $p > \sum a_i$		
A, A^{-1}		
$A \cdot (a_1, a_2, \ldots, a_n)$ mod p		
$(b_1, \ldots, b_n)$	$(b_1, \ldots, b_n)$	$(b_1, \ldots, b_n)$
		ASCII code
		plaintext – $(x_1, x_2, \ldots, x_n)$
C	C	$C = x_1 b_1 + x_2 b_2 + \ldots + x_n b_n$
$M = A^{-1} \cdot C$ mod p		
decompose M in terms of $(a_1, a_2, \ldots, a_n)$		

Example. Alice picks the super increasing sequence

$$(a_1, a_2, \ldots, a_8) = (2,\ 5,\ 9,\ 22,\ 47,\ 99,\ 203,\ 409)$$

the prime $p = 997$ and the encryption factor $A = 60$.

a. Compute the sequence b_i's that Alice will publish

b. Suppose Bob wants to send the letter "**b**" (ASCII code: 01100010) to Alice. Find his ciphertext C.

c. Suppose Alice receives $C = 1255$ from Bob. Decrypt this message.