CIS 3362 9/2/26

Vigenere Cipher

Battiste Alberti - Italian, he designed Trevi Fountain

Key - 2 substitution alphabets

HELLO

sub1 sub2 sub1 sub2 sub1

$\sim$ ~~18~~1586 Vigenere

Plaintext: $p_0\ p_1\ p_2 \cdots\ p_{n-1}$ (length n)

Keyword $k_0\ k_1\ k_2 \cdots k_{m-1}$ (length m)

$$c_i = (p_i + \underline{\underline{k_{i \% m}}}) \bmod 26$$

$$p_i = (c_i - k_{i \% m} + 26) \% 26$$

for code

Secure until early 1800s.

Paradigm Shift

Tools to Break Vigenere

1) Kasiski Test (discovered by Charles Babbage also)

2) Index of Coincidence

GOAL IS TO IDENTIFY THE KEYWORD LENGTH

3) Mutual Index of Coincidence

	i			j
Plain	THE		THE	THE
	$k_2 k_3 k_3$	....	$k_3 k_4 k_5$	$k_2 k_3 k_4$
	MQA			MQA

If repeated plaintext AND happens to line up cyclically, then you'll see repeated ciphertext

If repeated, there's a good chance plain is same + we've found a cycle length.

Let m = keyword length, then $m \mid (j-i)$

"$a \mid b$" - "b is divisible by a"

"There exists an integer c such that $b = ac$."

for all pairs of repeats (3 or more letters)

Call these differences in index

$d_1, d_2, d_3 \ldots \rightarrow m \mid d_1, m \mid d_2, \ldots$ 8

$m \mid 80, \; m \mid 96, \; m \mid 200 \Rightarrow m \mid \gcd(80, 96, 200)$

Index of Coincidence

Bag: 20 Snickers
10 Twix
50 Skittles
20 Hersheys

If you select 2 candies from the bag, what is the probability both are the same type?

$$p(2\,Sn) + p(2\,Tw) + p(2\,Sk) + p(2\,He)$$

$$\frac{20}{100} \times \frac{19}{99} + \frac{10 \times 9}{100 \times 99} + \frac{50 \times 49}{100 \times 99} + \frac{20 \times 19}{100 \times 99}$$

Sample Space: # ways to choose 2 items out of 100 $\to \binom{100}{2}$

$$\frac{\binom{20}{2} + \binom{10}{2} + \binom{50}{2} + \binom{20}{2}}{\binom{100}{2}}$$

$$\frac{\sum_{i=1}^{k} f_i(f_i - 1)}{n(n-1)}, \quad \sum_{i=1}^{k} f_i = n$$

k = # of types of object
f_i = # of copies of object i
n = total # of objects

$$\frac{2\times 20\times 19 + 90 + 50\times 49}{100 \times 99}$$

$$= \frac{4\times 19 + 9 + 5\times 49}{990}$$

$$\begin{array}{r} 85 \\ 245 \\ \hline 330 \end{array}$$

$$= \frac{76 + 9 + 245}{990}$$

$$= \frac{330}{990}$$

$$= \boxed{\frac{1}{3}}$$

$\boxed{\text{Index of Coincidence IS invariant under } \cancel{\text{and}} \text{ any Substitution}}$

IC (english text) $\sim .0676$

IC (random letters) $\sim \frac{1}{26} < .04$

Try $m = 2, 3, 4, \ldots$

$c_0 \; c_1 \; c_2 \; c_3 \; \ldots$

PUT LETTERS IN m bins:

bin 0: $c_0, c_m, c_{2m}, \ldots$

bin 1: c_1, c_{m+1}, c_{2m+1}

$\vdots$

bin $m-1$: $c_{m-1}, c_{2m-1}, c_{3m-1}, \ldots$

If m is correct $IC(b_0), IC(b_1), \ldots \sim .07$
but if m is wrong most will be lower

Once you get keyword length, try shifting bins to "look like" English. Plug those in + see if you make out any words

$K_2 K_3$ $K_7 K_9$ correct...

+ Do some guesswork to fill in letters that you might have guessed wrong

MUTUAL INDEX OF COINCIDENCE

Given multisets A and B, what's the probability 1 randomly selected item from A equals 1 randomly selected item from B?

$A = 10a, 20b, 40c$ 70

shift −1 ← $B = 50a, 15b, 25c$ 90

$B' = 15b, 25c, 50a$

$$\frac{10}{70} \times \frac{50}{90} + \frac{20}{70} \times \frac{15}{90} + \frac{40}{70} \times \frac{25}{90}$$

$$MIC(A, B') = \frac{10 \times 15 + 20 \times 25 + 40 \times 50}{70 \times 90}$$

```
int res = 0; ans = 0;
for (int i = 0; i < 26; i++ )
    int tmp = computeMIC (english, f );
    if (tmp > ans) {
        ans = tmp;
        res = i
    }
    f = cyclicLeftShift (f);
}
// value of res is most-likely
// shift char for this bin.
```

Compute 2 or 3 highest MICs and try all 2^m or 3^m combos looking for words ~~for~~ from a dictionary OR a few common words