The Vigenére Cipher, Matrix Cipher, Issues, and One-Time Pad

Lecture 03

The Vigenère cipher

Key: A word or phrase. Example: *dog* = *(3,14,6)*.
Easy to remember and transmit.
Example using *dog*.
Shift 1st letter by 3
Shift 2nd letter by 14
Shift 3nd letter by 6
Shift 4th letter by 3
Shift 5th letter by 14
Shift 6th letter by 6, etc.

Jacob Prinz is a Physics Major
Jacob Prinz isaPh ysics Major

encrypts to

MOIRP VUWTC WYDDN BGOFG SDXUU

The Vigenère cipher

Key: $k = (k_1, k_2, \ldots, k_n)$.
Encrypt (all arithmetic is mod 26)

$$Enc(m_1, m_2, \ldots, m_N) =$$

$$m_1 + k_1, m_2 + k_2, \ldots, m_n + k_n,$$

$$m_{n+1} + k_1, m_{n+2} + k_2, \ldots, m_{n+n} + k_n,$$

$$\ldots$$

Decrypt Decryption just reverse the process

The Vigenère cipher

- Size of key space?
  - If keys are 14-char then key space size $26^{14} \approx 2^{66}$
  - If variable length keys, even more.
  - Brute-force search infeasible
- Is the Vigenère cipher secure?
- Believed secure for many years...
- Might not have even been secure then...

Cracking Vig cipher: Step One-find Keylength

Assume T is a text encoded by Vig, key length L unknown.
For $0 \le i \le L-1$, letters in pos $\equiv i \pmod{26}$ – same shift.
Look for a sequence of (say) 3-letters to appear (say) 4 times.

Example: aiq appears in the
57-58-59th slot, 87-88-89th slot 102-103-104th slot
162-163-164th slot

Important: Very likely that aiq encrypted the same 3-letter sequence and hence the length of the key is a divisor of
87-57=30 102-87=15 162-102=60
The only possible L's are 1,3,5,15.

Good Enough: We got the key length down to a small finite set.

Important Point about letter Freq

Assume (and its roughly true): In an English text of length N:

e occurs $\sim 13\%$ t occurs $\sim 9\%$ a occurs $\sim 8\%$

Etc- other letters have frequencies that are true for all texts.

Important Point about letter Freq

Assume (and its roughly true): In an English text of length N:

e occurs $\sim 13\%$ t occurs $\sim 9\%$ a occurs $\sim 8\%$

Etc- other letters have frequencies that are true for all texts.

Assume (and its roughly true): In an English text of length N, if $i \ll N$, then if you take every ith letter of T:

e occurs $\sim 13\%$ t occurs $\sim 9\%$ a occurs $\sim 8\%$

Etc- other letters same frequencies as normal texts.

Important Point about letter Freq

Assume (and its roughly true): In an English text of length N:

e occurs $\sim 13\%$ t occurs $\sim 9\%$ a occurs $\sim 8\%$

Etc- other letters have frequencies that are true for all texts.

Assume (and its roughly true): In an English text of length N, if $i \ll N$, then if you take every ith letter of T:

e occurs $\sim 13\%$ t occurs $\sim 9\%$ a occurs $\sim 8\%$

Etc- other letters same frequencies as normal texts.

Relevant to us:

$\vec{q}$ freq of every Lth letter: then $\sum_{i=1}^{26} q_i^2 \approx 0.065$.

$\vec{q}$ is NOT (we won't define that rigorously): $\sum_{i=1}^{26} q_i^2$ MUCH lower.

Cracking Vig cipher: Step One-find Keylength

Let K be the set of possible key lengths. K is small. For every $L \in K$:

- Form a stream of every Lth character.
- Find the frequencies of that stream: $\vec{q}$.
- Compute $Q = \sum q_i^2$
- If $Q \approx 0.065$ then YES L is key length.
- If Q much less than 0.065 then NO L is not key length.
- One of these two will happen
- Just to make sure, check another stream.

Note: Differs from Is English:

Is English wanted to know if the text was actually English

What we do above is see if the text has same dist of English, but okay if diff letters. E.g., if z is 13%, a is 9%, and other letters have roughly same numbers as English then we know the stream is SOME Shift. We later use Is English to see which shift.

A Note on Finding Keylength

We presented Method ONE:

1. Find phrase of length x appearing y times. Differences D.
2. K is set of divisors of all $L \in D$. Correct keylength in K.
3. Test $L \in K$ for key length until find one that works.

Or could try all key lengths up to a certain length, Method TWO:

1. Let $K = \{1, \ldots, 100\}$ (I am assuming key length ≤ 100).
2. Test $L \in K$ for key length until find one that works.

Note: With modern computers use Method TWO. In days of old eyeballing it made Method ONE reasonable.

Cracking the Vig cipher: Step Two-Freq Anal

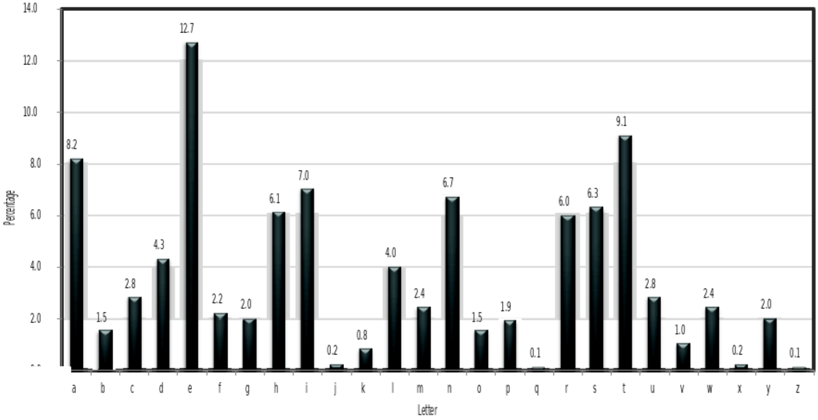
After Step One we have the key length L. Note:

- Every L^{th} character is "encrypted" using the same shift.
- Important: Letter Freq still hold if you look at every L 14th letter!

Step Two:

1. Separate text T into L streams depending on position mod L
2. For each steam try every shift and use Is English to determine which shift is correct.
3. You now know all shifts for all positions. Decrypt!

Using plaintext letter frequencies



Gen 2-letter Sub and Matrix Codes

Lecture 03

Shift, Affine, Vig, Gen Sub, Easy to Crack

Shift, Affine, Vig all 1-letter substitutions. Freq cracked them.
Idea: Lets substitute two letters at a time.
An Idea Which History Passed By:
Definition: Gen Sub 2-Cipher with perm f on $\{0, \ldots, 25\}^2$.

1. Encrypt via $xy \rightarrow f(xy)$.
2. Decrypt via $xy \rightarrow f^{-1}(xy)$

Why never used?

1. It was used but they kept it hidden and still not known!
2. The key length is roughly $26^2 \times 10 = 6760$ bits.
3. Old days: hard to use. Now: easy to crack.

Need bijection of $\{0, \ldots, 25\} \times \{0, \ldots, 25\}$ that is easy to use.

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \to M(xy)$.
2. Decrypt via $xy \to M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \to M(xy)$.
2. Decrypt via $xy \to M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.

HEY- WAIT A MINUTE!

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \rightarrow M(xy)$.
2. Decrypt via $xy \rightarrow M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.
HEY- WAIT A MINUTE!
Easy to see if M^{-1} exists? Easy to find M^{-1}?

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \to M(xy)$.
2. Decrypt via $xy \to M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.
HEY- WAIT A MINUTE!
Easy to see if M^{-1} exists? Easy to find M^{-1}?
Is Bill punking you ... again?

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \rightarrow M(xy)$.
2. Decrypt via $xy \rightarrow M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.

HEY- WAIT A MINUTE!

Easy to see if M^{-1} exists? Easy to find M^{-1}?

Is Bill punking you ... again? No he is not.

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then

$$M^{-1} = \frac{1}{ad - bc} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Do you recognize the expression $ad - bc$?

The Matrix Cipher

Definition: Matrix Cipher. Pick M a 2×2 matrix.

1. Encrypt via $xy \rightarrow M(xy)$.
2. Decrypt via $xy \rightarrow M^{-1}(xy)$

Encode: Break T into blocks of 2, apply M to each pair.

Decode: Do the same only with M^{-1}.
HEY- WAIT A MINUTE!
Easy to see if M^{-1} exists? Easy to find M^{-1}?
Is Bill punking you ... again? No he is not.

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then

$$M^{-1} = \frac{1}{ad - bc} \times \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Do you recognize the expression $ad - bc$? Determinant!

Inverse Matrix in $\mathbb{C}$ and in Mods

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

1. Matrix M over $\mathbb{C}$ has an inverse iff $ad - bd \neq 0$.
2. Matrix M over Mod n has an inverse iff $ad - bc$ is rel prime to n iff $ad - bc$ has an inverse in Mod n.
3. Matrix M over Mod 26 has an inverse iff $ad - bc$ is rel prime to 26 iff $ad - bc$ has no factors of 2 or 13 iff has an inverse in Mod 26.

Stuff to know for Special Lecture on Sept 24:

1. A matrix is invertible iff all of the rows are linearly ind.
2. If over $\mathbb{Z}_p$ where p is a prime then more like $\mathcal{C}$- all numbers have inverses so need $ad - bc \neq 0$.

The Matrix Cipher

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Good News:

1. Can test if M^{-1} exists, and is so find it, easily.
2. M small, so Key small.
3. Applying M or M^{-1} to a vector is easy computationally.

Bad News:

1. Eve CAN crack using frequencies of pairs of letters.
2. Eve CAN crack – Key space has $< 26^4 = 456976$. Small.

So what to do?

The Matrix Cipher

Definition: Matrix Cipher. Pick n and M an $n \times n$ matrix with det rel prime to 26.

1. Encrypt via $\vec{x} \rightarrow M(\vec{x})$.
2. Decrypt via $\vec{y} \rightarrow M^{-1}(\vec{y})$

We'll take $n = 30$.

The Matrix Cipher

Definition: Matrix Cipher. Pick n and M an $n \times n$ matrix with det rel prime to 26.

1. Encrypt via $\vec{x} \to M(\vec{x})$.
2. Decrypt via $\vec{y} \to M^{-1}(\vec{y})$

We'll take $n = 30$.

1. Can determine if M has inv and if so find it easily.
2. M still small, so Key small.
3. Applying M or M^{-1} to a vector is easy computationally.
4. Eve can crack using freqs of 30-long sets of letters? Hard?
5. Eve cannot use brute force – Key Space is $\sim 26^{900}$.

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

1. If Eve just has ciphertext then brute force needs of 26^{n^2} possibilities. Can get that down to 26^n.

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

1. If Eve just has ciphertext then brute force needs of 26^{n^2} possibilities. Can get that down to 26^n.
2. 26^n is still large. Can Eve do better?

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

1. If Eve just has ciphertext then brute force needs of 26^{n^2} possibilities. Can get that down to 26^n.
2. 26^n is still large. Can Eve do better?
   Seems to be Unknown to Science!

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

1. If Eve just has ciphertext then brute force needs of 26^{n^2} possibilities. Can get that down to 26^n.
2. 26^n is still large. Can Eve do better?
   Seems to be Unknown to Science!
   So why is it not used? Discuss!

Is Matrix Cipher Uncrackable?

VOTE: Yes, No, Unknown to Science, Other.

1. If Eve just has ciphertext then brute force needs of 26^{n^2} possibilities. Can get that down to 26^n.
2. 26^n is still large. Can Eve do better?
   Seems to be Unknown to Science!
   So why is it not used? Discuss!
3. In reality Eve has prior messages and what they coded to, so from that she can easily crack it. (Next Slide.) That is why not used.

Cracking Matrix Cipher

Example using 2×2 Matrix Cipher.
Eve learns that (19,8) encrypts to $(3, 9)$. Hence:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 19 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix}$$

So

$$\begin{aligned} 19a + 8b &= 3 \\ 19c + 8d &= 9 \end{aligned}$$

Two linear equations, Four variables

Cracking Matrix Cipher

Example using 2×2 Matrix Cipher.
Eve learns that (19,8) encrypts to $(3, 9)$. Hence:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 19 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix}$$

So

$$\begin{aligned} 19a + 8b &= 3 \\ 19c + 8d &= 9 \end{aligned}$$

Two linear equations, Four variables

If Eve learns one more 2-letter message decoding then she will have

Four linear equations, Four variables

which she can solve! Yeah? Boo? Depends whose side you are on.