

Monoalphabetic ciphers

Chris Brew

January 9, 2008

Codes and Ciphers

Shift ciphers

Things that can go wrong?

Making things harder for the cryptanalyst

Exhaustive enumeration: 2

CPZPAPUNHBUAZJHUILHWYVISLT
 DQAQBQVOICVBAKIVJMIXZWJTMU
 ERBRCRWPJDWCBLJWKNJYAXKUNV

...

UHRHSHMFZTMSRBZMADZOQNAKDL
 VISITINGAUNTSKANBEAPROBLEM
 WJTJUJOHBVOUTDBOCFBQSPCMFN
 XKUKVKPICWPVUECPDGCRTQDNGO
 YLVLWLQJDXQWVFDQEHSUREOHP
 ZMWMXMRKEYRXWGERFIETVSFPIQ
 ANXNYNSLFZSYXHFSGJFUWTGQJR
 BOYOZOTMGATZYIGTHKGVXUHRKS

<===

What if it is not a shift cipher?

- ▶ What if you assume it is a shift cipher, but it actually isn't.

What if it is not a shift cipher?

- ▶ What if you assume it is a shift cipher, but it actually isn't.
- ▶ None of the rows will contain English.

What else?

- ▶ What if more than one of the rows contain English?

What else?

- ▶ What if more than one of the rows contain English?
- ▶ Then the cryptanalyst can't tell which of the messages are intended. This is because they do not know how much the alphabet is shifted.

What else?

- ▶ What if more than one of the rows contain English?
- ▶ Then the cryptanalyst can't tell which of the messages are intended. This is because they do not know how much the alphabet is shifted.
- ▶ The authorised recipient does know which message is intended, because they do know how much the alphabet is shifted.

Keys

- ▶ Encryption involves a **cryptosystem** (in this case just the shift cipher).

Encryption and Decryption Keys

CPZPAPUNHBUAZJHUILHWYVISLT	0	0
DQAQBQVOICVBAKIVJMIXZWJTMU	1	25
ERBRCRWPDWCBLJWKNJYAXKUNV	2	24
...		
UHRHSHMFZTMSRBZMADZOQNAKDL	18	8
VISITINGAUNTSCANBEAPROBLEM	19	7
WJTJUJOHBVOUTDBOCFBQSPCMFN	20	6
XKUKVKPICWPVUECPDGCRTQDNGO	21	5
YLVWLQJDXQWVFDQEHSUREOHP	22	4
ZMWMXMRKEYRXWGERFIETVSFPIQ	23	3
ANXNYNSLFSZYXHFSGJFUWTGQJR	24	2
BOYOZOTMGATZYIGTHKGVXUHRKS	25	1

The left column is the encryption key and the right column is

Exhaustive Enumeration

- ▶ The first reason why the exhaustive strategy works is that there are only very few keys

Exhaustive Enumeration

- ▶ The first reason why the exhaustive strategy works is that there are only very few keys
- ▶ One way to defeat this strategy is to vastly increase the number of available keys.

Exhaustive Enumeration

- ▶ The first reason why the exhaustive strategy works is that there are only very few keys
- ▶ One way to defeat this strategy is to vastly increase the number of available keys.
- ▶ The second reason why the exhaustive strategy works is that English is highly recognizable as such.

Exhaustive Enumeration

- ▶ The first reason why the exhaustive strategy works is that there are only very few keys
- ▶ One way to defeat this strategy is to vastly increase the number of available keys.
- ▶ The second reason why the exhaustive strategy works is that English is highly recognizable as such.
- ▶ One way to defeat this is to deliberately break up the patterns of ENGLISH spelling.

Increasing the number of keys

- ▶ Instead of using a shifted alphabet, allow the use of any shuffle of the alphabet.

Increasing the number of keys

- ▶ Instead of using a shifted alphabet, allow the use of any shuffle of the alphabet.
- ▶ This is known as a **general monoalphabetic cipher**

Encryption and decryption keys

For encryption, look up letter in top line, write down equivalent from bottom line.

ABCDEFGHIJKLMN OPQRSTUVWXYZ
UFKCOQRGMYTHZEJBILDVPSWANX

For decryption, look up letter in bottom line, write down equivalent from top line.

XPDSNBHLQOCRIYEUFVKATWZJM
ABCDEFGHIJKLMN OPQRSTUVWXYZ

The only difference is that the encryption key has the columns sorted in order of their plaintext letter, but the decryption key has them in order of their cryptotext letter.

Increasing the number of keys

- ▶ A general monoalphabetic cipher clearly has a large number of possible keys.

Increasing the number of keys

- ▶ A general monoalphabetic cipher clearly has a large number of possible keys.
- ▶ Exactly how many possibilities are there?

How many random shuffles?

- ▶ We could put any of 26 letters in first position

How many random shuffles?

- ▶ We could put any of 26 letters in first position
- ▶ Any of 25 remaining letters in position two.

How many random shuffles?

- ▶ We could put any of 26 letters in first position
- ▶ Any of 25 remaining letters in position two.
- ▶ And so on . . .

How many random shuffles?

- ▶ We could put any of 26 letters in first position
- ▶ Any of 25 remaining letters in position two.
- ▶ And so on ...
- ▶ There are
 $26 \times 25 \times \dots \times 1 = 403,291,461,126,605,635,584,000,000$
possibilities

A paradox

- ▶ It would take trillions of years to blindly explore all these possibilities.

A paradox

- ▶ It would take trillions of years to blindly explore all these possibilities.
- ▶ Yet breaking a monoalphabetic cipher is childs play.

A paradox

- ▶ It would take trillions of years to blindly explore all these possibilities.
- ▶ Yet breaking a monoalphabetic cipher is childs play.
- ▶ How can this be?

A paradox

- ▶ It would take trillions of years to blindly explore all these possibilities.
- ▶ Yet breaking a monoalphabetic cipher is childs play.
- ▶ How can this be?
- ▶ Obviously, the answer does not involve blind search.

Invariance

- ▶ The key observation is that a monoalphabetic cipher always maps plaintext A to the same cryptotext letter.

Invariance

- ▶ The key observation is that a monoalphabetic cipher always maps plaintext A to the same cryptotext letter.
- ▶ We don't know what this letter is, but we know there will be one.

Invariance

- ▶ The key observation is that a monoalphabetic cipher always maps plaintext A to the same cryptotext letter.
- ▶ We don't know what this letter is, but we know there will be one.
- ▶ Every other plaintext letter has its own unique partner in the cryptotext

Invariance

- ▶ The key observation is that a monoalphabetic cipher always maps plaintext A to the same cryptotext letter.
- ▶ We don't know what this letter is, but we know there will be one.
- ▶ Every other plaintext letter has its own unique partner in the cryptotext
- ▶ This is called an **invariant** of the cryptosystem, and guarantees that there will be orderly patterns of various sorts in the cryptotext.

Invariance

- ▶ The key observation is that a monoalphabetic cipher always maps plaintext A to the same cryptotext letter.
- ▶ We don't know what this letter is, but we know there will be one.
- ▶ Every other plaintext letter has its own unique partner in the cryptotext
- ▶ This is called an **invariant** of the cryptosystem, and guarantees that there will be orderly patterns of various sorts in the cryptotext.
- ▶ Cryptology is all about invariants. Cryptanalysts try hard to uncover and use them. Code makers try hard to avoid and/or obscure them.

Repetition patterns

- ▶ Plaintext: OHIO STATE BUCKEYES
- ▶ One-shift: PIJP TUBUF CVDJFZFP
- ▶ Caesar: RKLR VWDWH EXFLHBHV
- ▶ Random: WV BW FMLMS JTYQSKSF

Repetition patterns

O H I O S T A T E B U C K E Y E S
P I J P T U B U F C V D J F Z F P
R K L R V W D W H E X F L H B H V
W V B W F M L M S J T Y Q S K S F
0 1 2 0 4 5 6 5 8 9 10 11 12 8 14 8 4