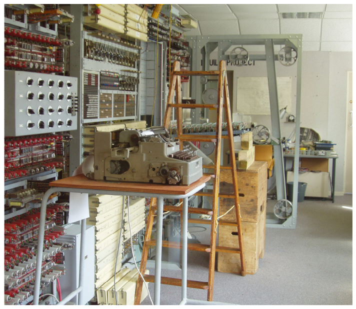
Class 2: The Lorenz Cipher and the Postman's Computer

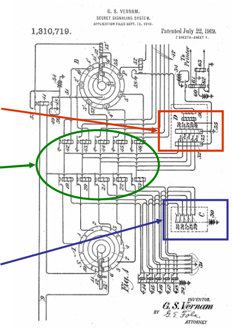


Colossus Rebuilt, Bletchley Park, Summer 2004

http://www.cs.virginia.edu/jillcrypto

One-Time Pad

Vernam [1917]
(AT&T Bell Labs)



The Baudot Code
(like Morse Code, not a cipher)

A 00011	H 10100		*space* 00100
B 11001	I 00110	… …	*return* 01000
C 01110	J 01011	V 11110	*line feed* 00010
D 01001	K 01111	W 10011	*letter shift* 11111
E 00001	L 10010	X 11101	*figure shift* 11011
F 01101	M 11100	Y 10101	*error* 00000
G 11010	N 01100	Z 10001	

Encode 32 letters using 5 on/off signals

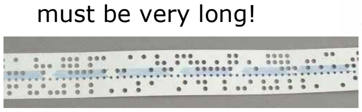
Why perfectly secure?

For any given ciphertext, all plaintexts are equally possible.

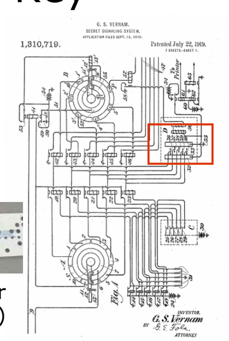
Ciphertext: **J** = 01001
Key1: **I** = 00110
Plaintext1: 01111 = **K**
Key2: **L** = 10010
Plaintext2: = **11011 = shift**

Vernam's Key

- A long paper tape with random letters on it (using Baudot code)
- Cannot reuse key – tape must be very long!



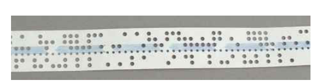
This has 6 holes per letter (not Baudot code)



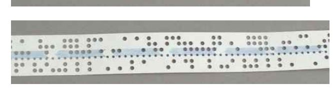
Morehouse's Improvement

- Like Vernam machine, but with two key tapes

Tape 1
(999 letters)



Tape 2
(1000 letters)



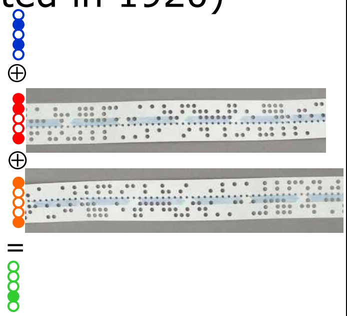
Morehouse's Improvement (patented in 1920)

Message $\oplus$

Tape 1
(999 letters) $\oplus$

Tape 2
(1000 letters) =

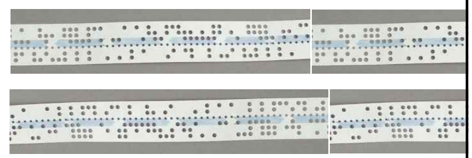
Ciphertext



Looping Tapes

Tape 1
(999 letters)

Tape 2
(1000 letters)



The tape equivalent to Tape 1 $\oplus$ Tape 2 would not repeat for 999 * 1000 letters!

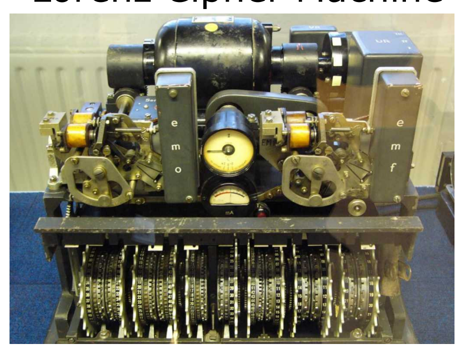
Note: it is no longer a perfect cipher though. Some keys are not possible after 1001 letters.

Lorenz Cipher

- Based on the Vernam and Morehouse
  - Used Baudot code
- Believed managing long paper tapes during wartime was too difficult
- Machine generates key sequence
  - If two machines start in same configuration, same key sequence
  - Will not repeat for ~ 10^{19} letters

All words ever spoken or written by all humans is estimated around 10^{18} letters

Lorenz Cipher Machine

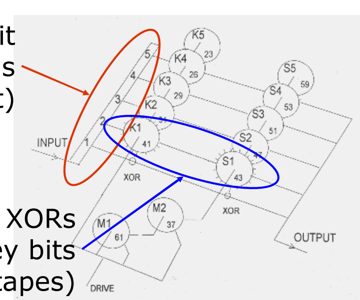


Lorenz Wheels

12 wheels
501 pins
total (set
to control
wheels)

Wheel Operation

Bit channels (5 for Baudot)

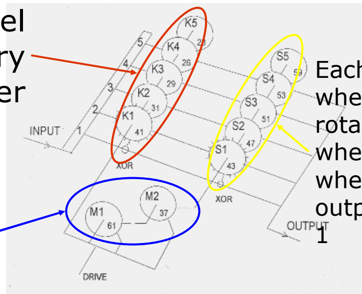


Two XORs with key bits (like paper tapes)

Wheel Operation

Each K wheel rotates every letter

M wheels control if S wheels rotate



Each S wheel rotates when M wheels output 1

Use by Nazis

- Considered most secure cipher machine
- Messages between Hitler's army headquarters and European capital headquarters
- Each link had a slightly different system (British named them for fish):
  - Tunny: Vienna - Athens
  - Jelly: Berlin – Paris

Breaking Fish

- GCHQ learned about first Fish link (Tunny) in May 1941
  - Intercepted unencrypted Baudot-encoded test messages
- August 30, 1941: Big Break!
  - Operator retransmits failed message with same starting configuration
  - Gets lazy and uses some abbreviations, makes some mistakes
    - SPRUCHNUMMER/SPRUCHNR (Serial Number)

"Two Time" Pad

- Allies have intercepted:

  $C1 = M1 \oplus \mathbf{K1}$

  $C2 = M2 \oplus \mathbf{K1}$

  Same key used for both (same starting configuration)
- Breaking message:

$$C1 \oplus C2 = (M1 \oplus \mathbf{K1}) \oplus (M2 \oplus \mathbf{K1})$$
$$= (M1 \oplus M2) \oplus (\mathbf{K1} \oplus \mathbf{K1})$$
$$= M1 \oplus M2$$

"Cribs"

- Know: C1, C2 (intercepted ciphertext)

  $C1 \oplus C2 = M1 \oplus M2$
- Don't know M1 or M2
  - But, can make some guesses (cribs)
    - SPRUCHNUMMER
    - Sometimes allies moved ships, sent out bombers to help the cryptographers get good cribs
- Given guess for M1, calculate M2

  $M2 = C1 \oplus C2 \oplus M1$
- Once guesses that work for M1 and

Finding K1

- From the 2 intercepted messages, Col. John Tiltman worked on guessing cribs to find M1 and M2
  - 4000 letter message, found 4000 letter key
- Bill Tutte (recent Chemistry graduate) given task of determining machine structure from key
  - Already knew it was 2 sets of 5 wheels and 2 wheels of unknown function

Reverse Engineering Lorenz

- Looked at patterns of bits in key
- Found repeating sequence:
  - Repetition period of 41, learned first wheel had 41 pins
  - Similar for other wheels, determining S/M/K wheel structure
- After 6 months of hard work: determined likely machine structure that would generate K1

Intercepting Traffic

- Set up listening post to intercept traffic from 12 Lorenz (Fish) links
  - Different links between conquered capitals
  - Slightly different coding procedures, and different configurations
- 600 people worked on intercepting traffic
- Sent intercepts to Bletchley (usually by motorcycle courier)

Breaking Traffic

- Knew machine structure, but a different initial configuration was used for each message
- Need to determine wheel setting:
  - Initial position of each of the 12 wheels
  - 1271 possible starting positions
  - Needed to try them fast enough to decrypt message while it was still strategically valuable

Recognizing a Good Guess

- Intercepted Message (divided into 5 channels for each Baudot code bit)

  $Z_c = z_0 z_1 z_2 z_3 z_4 z_5 z_6 z_7 \ldots$

  $z_{c,i} = m_{c,i} \oplus x_{c,i} \oplus s_{c,i}$

  Message Key (parts from S-wheels and rest)
- Look for statistical properties
  - How many of the $z_{c,i}$'s are 0? ½ (not useful)
  - How many of $(z_{c,i+1} \oplus z_{c,i})$ are 0? ½

Double Delta

- $\Delta Z_{c,i} = Z_{c,i} \oplus Z_{c,i+1}$
- Combine two channels:

  $\Delta Z_{1,i} \oplus \Delta Z_{2,I} =$

  $\Delta M_{1,i} \oplus \Delta M_{2,i}$ > ½ Yippee!

  $\oplus \Delta X_{1,i} \oplus \Delta X_{2,i}$ = ½ (key)

  $\oplus \Delta S_{1,i} \oplus \Delta S_{2,i}$ > ½

Double Delta

$\Delta M_{1,i} \oplus \Delta M_{2,i}$ > ½ Yippee!

$\oplus \Delta X_{1,i} \oplus \Delta X_{2,i}$ = ½ (key)

$\oplus \Delta S_{1,i} \oplus \Delta S_{2,i}$ > ½

Why is $\Delta M_{1,i} \oplus \Delta M_{2,i} > ½$

Message is in German, more likely following letter is a repetition than random

Why is $\Delta S_{1,i} \oplus \Delta S_{2,i} > ½$

S-wheels only turn some of the time (when M-wheel is 1)

Actual Advantage

- Probability of repeating letters

  Prob[$\Delta M_{1,i} \oplus \Delta M_{2,i} = 0$] ~ 0.614

  3.3% of German digraphs are repeating
- Probability of repeating S-keys

  Prob[$\Delta S_{1,i} \oplus \Delta S_{2,i} = 0$] ~ 0.73

Prob[$\Delta Z_{1,i} \oplus \Delta Z_{2,I} \oplus \Delta X_{1,i} \oplus \Delta X_{2,i} = 0$]

= 0.614 * 0.73 + (1-0.614) * (1-0.73)

Δ M and S are 0 $\quad\Delta$ M and S are 1

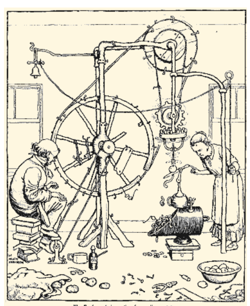
= 0.55

Using the Advantage

- If the guess of **X** is correct, should see higher than ½ of the double deltas are 0
- Try guessing different configurations to find highest number of 0 double deltas
- Problem:
  - \# of double delta operations to try one config
  - = length of Z * length of X
  - = for 10,000 letter message = 12 M for each setting * 7 ⊕ per double delta
  - = 89 M ⊕ operations

Heath Robinson

- Dec 1942: Decide to build a machine to do these ⊕s quickly, due June 1943
- Apr 1943: first Heath Robinson machine is delivered!
- Intercepted ciphertext on

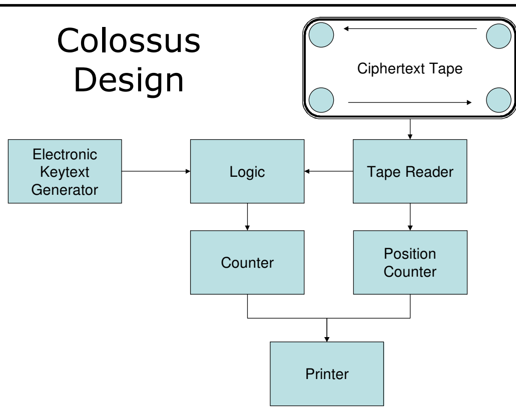


Heath Robinson, British Cartoonist (1872-1944)

Colossus

- Heath Robinson machines were too slow
- Colossus designed and first built in Jan 1944
- Replaced keytext tape loop with electronic keytext generator
- Speed up ciphertext tape:
  - 5,000 chars per second = 30 mph
  - Perform 5 double deltas simultaneously
  - Speedup = 2.5X for faster tape * 5X for parallelism

Colossus Design

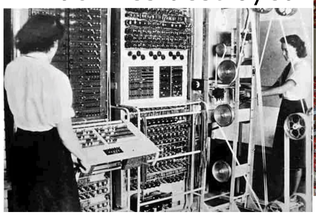


Impact on WWII

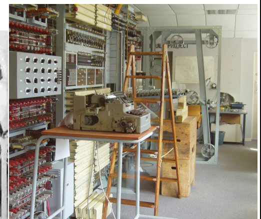
- 10 Colossus machines operated at Bletchley park
  - Various improvements in speed
- Decoded 63 million letters in Nazi command messages
- Learned German troop locations to plan D-Day (knew the deception was working)

Colossus History

- Kept secret after the war, all machines destroyed



During WWII



Rebuild, Bletchley Park, Summer 2004

Next Class

- Enigma and how it was broken
- Some similarities to Colossus:
  - Exploited operator errors
  - Built machines to quickly try possibilities