Cryptography in World War II
Jefferson Institute for Lifelong Learning at UVa
Spring 2006 **David Evans**

Class 3: Captain Ridley's Shooting Party

Turing's Hut 8 at Bletchley Park

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

Enigma

Enigma machine at Bletchley Park

- Invented commercially, 1923
- Used by German Navy, Army, Air Force
- About 50,000 in use
- Modified throughout WWII, believed to be perfectly secure
- Kahn's *Codebreakers* (1967) didn't know it was broken
- Turing's 1940 Treatise on Enigma declassified in 1996

Simple Substitution Ciphers (from Class 1)

```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
encrypt        decrypt
JIDKQACRSHLGWNFEXUZVTPMYOB
```

JILL ⇒ HSGG

Rotating Substitution Cipher

```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
encrypt
IDKQACRSHLGWNFEXUZVTPMYOBJID
```

J ⇒ H I ⇒ H L ⇒ N L ⇒ F

JILL ⇒ HHNF

Rotating Substitution Cipher

- Rotates the mapping every letter
  - Hides simple statistical properties of plaintext:
    - Frequency analysis defeated: E encrypts to different letters
    - Repeated letter will not encrypt the same way in different positions

Rotating Substitution Weaknesses

- Will repeat after 26 letters
  - If there is a lot of ciphertext, can still do frequency analysis on every 26th letter slides
- Some properties revealed
  - If we see repeated letters in ciphertext, what does it mean?

JILL ⇒ HHNF

Multiple Substitution Ciphers

ABCDEFGHIJKLMNO

JIDKQACRSHLGWNF

ABCDEFGHIJKLMNO

SQHLZNYKXUWVJRDFBETIMOGACP

J ⇒ K

This doesn't help at all: Any number of multiple simple substitutions can be replaced by one substitution!

Multiple Rotating Substitutions

Wheel 1: Rotate one position every letter

ABCDEFGHIJKLM

JIDKQACRSHLGW

Wheel 2: Rotate one position every 26 letters

ABCDEFGHIJKLM

SQHLZNYKXUWVJRDFBETIMOGACP

Now it only repeats when both wheels have cycled: 26*26 = 676 letters!

Multiple Rotating Substitutions

Wheel 1: Rotate one position every letter

ABCDEFGHIJKL

JIDKQACRSHLG

Wheel 2: Rotate one position every 26 letters

ABCDEFGHIJKLM

SQHLZNYKXUWVJ

Wheel 3: Rotate one position when wheel 2 cycles

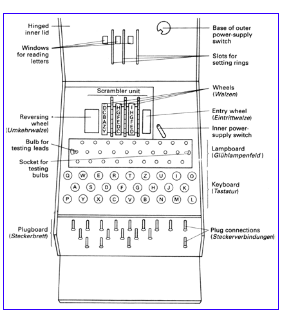
ABCDEFGHIJKLMNOPQRSTUVWXYZ

UAVGRDCBESYHLZOQKXTIMNJWFP

Now it only repeats when all 3 wheels have cycled: 26*26 * 26 = 17576 letters!

Enigma

Enigma Mechanics



- Three rotors (chosen from 5), scrambled letters
  - Each new letter, first rotor advances
  - Other rotors advance when ring is hit
- Reflector
- Plugboard

Rotor Wheel

Simple substitution

No letter maps to itself

Latch turns next rotor once per rotation

Settings

- Plugboard: swap pairs of letters
  - Number of plugs varied (≤ 6 until 1939, up to 10 after)
- Rotors
  - Before 1939 – Three rotors (choose order)
  - After – Choose 3 from set of 5 rotors
  - Orientations (3) – start orientations of the 3 rotors
  - Ring settings (2) – when next ring advances
- Reflector
  - Fixed symmetric substitution ($A \to B \Rightarrow B \to A$)
    Involution: if we do it twice, get original back

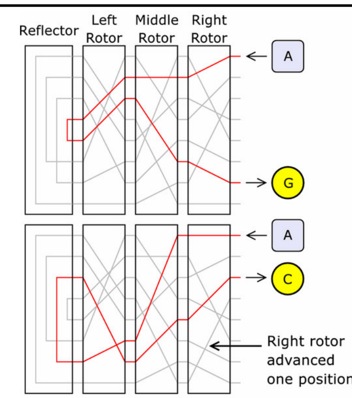
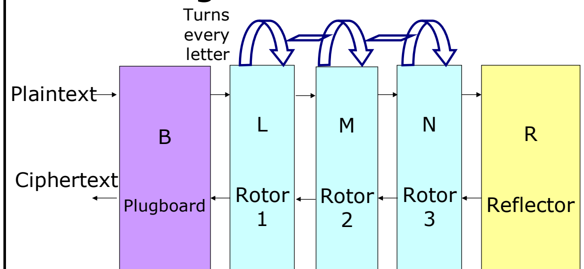


Image from http://en.wikipedia.org/wiki/Image:Enigma-action.png

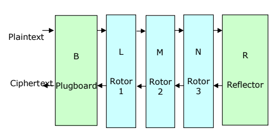
Three Rotor Wheels

Enigma Schematic



Ciphertext = $B^{-1}L^{-1}M^{-1}N^{-1}RNMLB$(Plaintext)

Does Decryption Work?

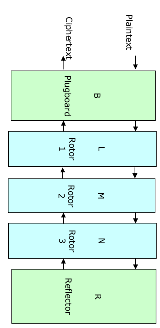


C = $B^{-1}L^{-1}M^{-1}N^{-1}RNMLB$(P)

P = $B^{-1}L^{-1}M^{-1}N^{-1}RNMLB$(C)

= $B^{-1}L^{-1}M^{-1}N^{-1}RNMLB(B^{-1}L^{-1}M^{-1}N^{-1}RNMLB(P))$

R is an involution
($A \to B \Rightarrow B \to A$)



Key Space

Plugboard with 6 plugs:
(26 * 25/2) * … * (16*15 / 2) / 6! ≈ 10^{11}

Rotors: $(26!)^3 \approx 4 * 10^{26}$

Ring settings: 26^2 = 676

Message Key: 26^3 = 17576

Reflector:
(26 * 25 / 2) * (24 * 23 / 2) * … * (2 * 2) / 13! ≈ 8 * 10^{12}

Total: ≈ 6 * 10^{110} (not all are different)
>> 10^{84} atoms in the universe

Reducing Key Space

Plugboard with 6 plugs ≈ 10^{11}

Rotors: $(26!)^3 \approx 4 * 10^{26}$

Ring settings: $26^2 = 676$

Message Key: $26^3 = 17576$

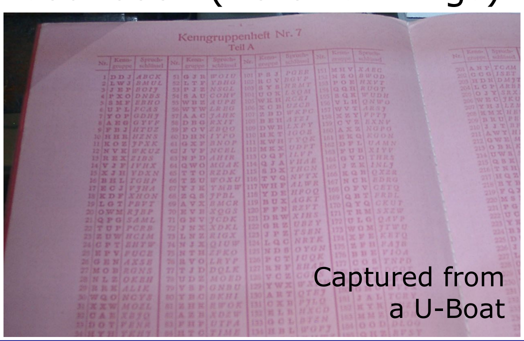
Reflector: $\approx 8 * 10^{12}$

Capture a Machine

"This fictional movie about a fictional U.S. submarine mission is followed by a mention in the end credits of those actual British missions. Oh, the British deciphered the Enigma code, too. Come to think of it, they pretty much did everything in real life that the Americans do in this movie."

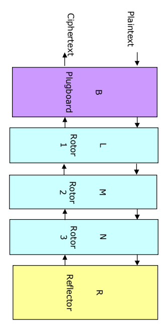
Roger Ebert's review of **U-571**

Codebook (Rotor Settings)



Captured from
a U-Boat

Key Space



Plugboard with 6 plugs:
$(26 * 25/2) * … * (16*15 / 2) / 6! \approx 10^{11}$

Rotors: $(26!)^3 \approx 4 * 10^{26}$ 5 C 3 = 60

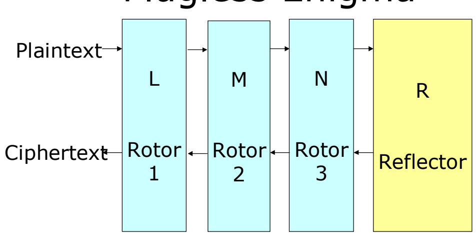
Ring settings: $26^2 = 676$

Message Key: $26^3 = 17576$

Reflector:
$(26 * 25 / 2) * (24 * 23 / 2) * … * (2 * 2) / 13! \approx 8 * 10^{12}$ 1

Total: $\approx 7 * 10^{19}$
($> 2^{64}$, still too big for exhaustive search

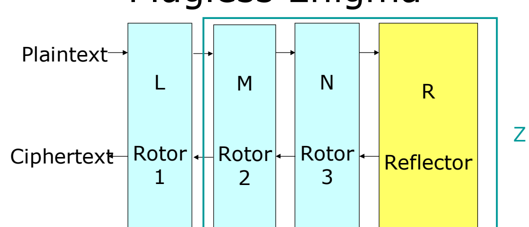
Plugless Enigma



$$C = L^{-1}M^{-1}N^{-1}RNML(P)$$

Used in Spanish Civil War (1937-9) by all participan
(including British, Germans and Spanish)

Plugless Enigma

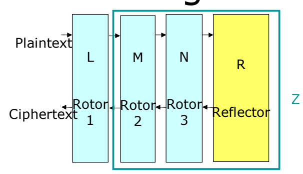


$C = L^{-1}ZL(P)$

$L(C) = ZL(P)$

Probable words (4-10 letters)
What is the probability that Rotor 2 and Rotor 3 do not move in 4 letter c
= 22/26 = .85

Plugless Enigma



$C = L^{-1}ZL(P)$

$L(C) = ZL(P)$

Z is a fixed substitution (monoalphabetic) if R2&3 don't

Guess a crib – have C and P_{guess}

$L(C) = ZL(P_{guess})$

Try possible rotors and starting positions for L:

3 rotor choices * 26 starting positions = 78

L_i = effect of Rotor 1 in the i^{th} rotation position

Batons Attack

C = XTSWVUINZ

P_{guess} = wehrmacht ("armed forces")

	ABCDEFGHIJKLMNOPQRSTUVWXYZ
L_1 (X) = Z L_1 (w)	EKMFLGDQVZNTOWYHXUSPAIBRCJ
L_2 (T) = Z L_2 (e)	JEKMFLGDQVZNTOWYHXUSPAIBRC
L_3 (S) = Z L_3 (h)	CJEKMFLGDQVZNTOWYHXUSPAIBR
L_4 (W) = Z L_4 (r)	RCJEKMFLGDQVZNTOWYHXUSPAIB
L_5 (V) = Z L_5 (m)	BRCJEKMFLGDQVZNTOWYHXUSPAI
L_6 (U) = Z L_6 (a)	IBRCJEKMFLGDQVZNTOWYHXUSPA
L_7 (I) = Z L_7 (c)	AIBRCJEKMFLGDQVZNTOWYHXUSP

For a given starting rotor setting, solve for Z

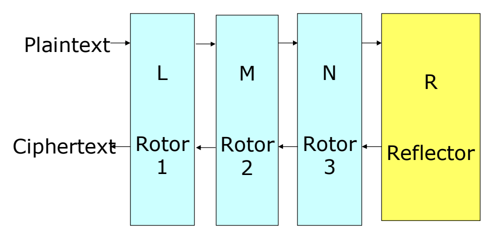
$\boxed{1: R = Z(B)}$ 2: S = Z(F) 3: X = Z(G) 4: P = Z(Y)

5: U = Z(V) 6: H = Z(I) $\boxed{7: M = Z(B)}$

Batons Attack

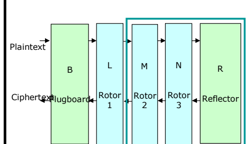
- We know Z is:
  - Function: contradiction if $Z(x) \neq Z(x)$
  - Involution: contradiction if $Z(x) = y$ & $Z(y) \neq x$
- Find a rotor setting with no contradictions
  - Long enough crib, there will only be one
  - But if crib is too long, need to deal with R2 moving
- List of probable 4-10 letter words
- Catalog to map Z to rotor settings for R2 and R3

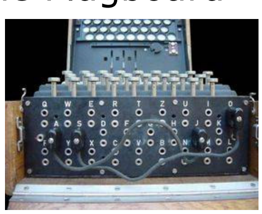
Plugless Enigma



Ideas for making Batons attack harder?

Enter the Plugboard



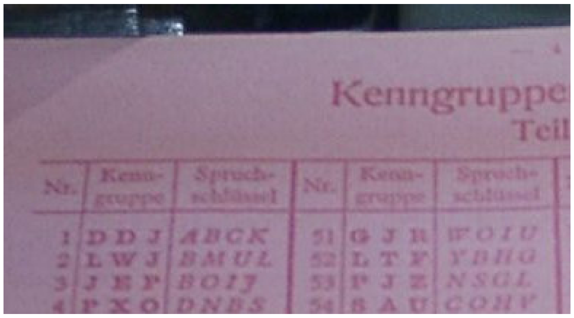


6 plugs: (26*25)/2 * (24*23)/2 * …
* (16*15/2) / 6!
~ 10^{11} times more keys

Operation

- Day key (distributed in code book)
- Each message begins with message key ("randomly" chosen by sender) encoded using day key
- Message key sent twice to check
- After receiving message key, re-orient rotors according to key

Codebook Zoom



Repeated Message Key

$P = P_1P_2P_3P_1P_2P_3$

$C_1 = E_1(P_1) = B^{-1}\mathbf{L_1}^{-1}M^{-1}N^{-1}RNM\mathbf{L_1}B(P_1)$
$C_4 = E_4(P_1) = B^{-1}\mathbf{L_4}^{-1}M^{-1}N^{-1}RNM\mathbf{L_4}B(P_1)$

$P_1 = E_1(C_1) = B^{-1}L_1^{-1}M^{-1}N^{-1}RNML_1B(C_1)$
$P_1 = E_4(C_4) = B^{-1}L_4^{-1}M^{-1}N^{-1}RNML_4B(C_4)$

$E_4 o E_1(C_1) = E_4(P_1) = C_4$
$E_4 o E_1 = B^{-1}L_1^{-1}M^{-1}N^{-1}RNML_1\boxed{B\,B^{-1}}L_4^{-1}M^{-1}N^{-1}RNML_4B$
$= B^{-1}L_1^{-1}M^{-1}N^{-1}RNML_1L_4^{-1}M^{-1}N^{-1}RNML_4B$

Letter Permutations

Symmetry of Enigma:

if $E_{pos}(x) = y$ we know $E_{pos}(y) = x$

Given message openings

DMQ **V**BM $E_1(m_1) = D$ $E_4(m_1) = V$
$E_1 o E_4(D) = V$
VON **P**UY => $E_1(D) = m_1$
PUC **F**MQ => $E_4(E_1(D)) = V$

With enough message openings, we can build complete cycles for each position pair:

$E_1 o E_4 =$ (DVPFKXGZYO) (EIJMUNQLHT) (BC) (RW) (A) (S)

Note: Cycles must come in pairs of equal length

Composing Involutions

- E_1 and E_2 are involutions ($x \rightarrow y \Rightarrow y \rightarrow x$)
- Without loss of generality, we can write:
  E_1 contains (a_1a_2) (a_3a_4) … $(a_{2k-1}a_{2k})$
  E_2 contains (a_2a_3) (a_4a_5) … $(a_{2k}a_1)$

E_1	E_2
$a_1 \leftrightarrow a_2$	$a_2 \leftrightarrow x = a_3$ or $x = a_1$
$a_3 \leftrightarrow a_4$	$a_4 \leftrightarrow x = a_5$ or $x = a_1$

Why can't x be a_2 or a_3?

Rejewski's Theorem

E_1 contains (a_1a_2) (a_3a_4) … $(a_{2k-1}a_{2k})$
E_4 contains (a_2a_3) (a_4a_5) … $(a_{2k}a_1)$

E_1E_4 contains $(a_1a_3a_5…a_{2k-1})$
$(a_{2k}a_{2k-2}… a_4a_2)$

- The composition of two involutions consists of pairs of cycles of the same length
- For cycles of length n, there are n possible factorizations

Factoring Permutations

$E_1E_4 =$ (DVPFKXGZYO) (EIJMUNQLHT) (BC) (RW) (A) (S)

(A) (S) = (AS) o (SA)
(BC) (RW) = (BR)(CW) o (BW)(CR)
or = (BW)(RC) o (WC) (BR)

(DVPFKXGZYO) (EIJMUNQLHT)
= (DE)(VI)… or (DI)(VJ) … or (DJ)(VM) … … (DT)(VE) 10 possibilities

How many factorizations?

(DVPFKXGZYO) (EIJMUNQLHT)

E_1	E_2
$D \leftrightarrow a_2$	$a_2 \leftrightarrow V$
$V \leftrightarrow a_4$	$a_4 \leftrightarrow P$

Once we guess a_2 everything else must follow!
So, only *n* possible factorizations for an *n*-letter cycle

Total to try = 2 * 10 = 20

E_2E_5 and E_3E_6 likely to have about 20 to try also

⇒ About 20^3 (8000) factorizations to try

(still too many in pre-computer days)

Luckily…

- Operators picked message keys ("cillies")
  - Identical letters
  - Easy to type (e.g., QWE)
- If we can guess $P_1 = P_2 = P_3$ (or known relationships) can reduce number of possible factorizations
- If we're lucky – this leads to $E_1 \ldots E_6$

Solving?

$E_1 = B^{-1}L^{-1}Q\,LB$

$E_2 = B^{-1}L^{-2}QL^2B$

$E_3 = B^{-1}L^{-3}QL^3B$

$E_4 = B^{-1}L^{-4}QL^4B$

$E_5 = B^{-1}L^{-5}QL^5B$

$E_6 = B^{-1}L^{-6}QL^6B$

6 equations, 3 unknowns

Not known to be efficiently solvable

Solving?

$E_1 = B^{-1}L^{-1}Q\,LB$

$BE_1B^{-1} = L^{-1}Q\,L$

Often, know plugboard settings (didn't change frequently)

6 equations, 2 unknowns – solvable

6 possible arrangements of 3 rotors, 26^3 starting locations
= 105,456 possibilities

Poles spent a year building a catalog of cycle structures covering all of them (until Nov 1937): 20 mins to break

Then Germans changed reflector and they had to start over.

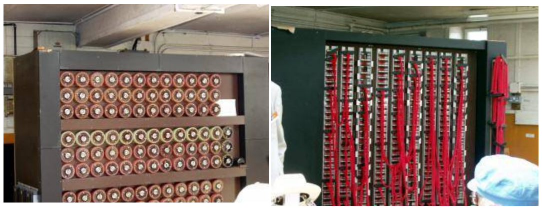
1939

- Early 1939 – Germany changes scramblers and adds extra plugboard cables, stop double-transmissions
  - Poland unable to cryptanalyze
- 25 July 1939 – Rejewski invites French and British cryptographers
  - Gives England replica Enigma machine constructed from plans, cryptanalysis
- 1 Sept 1939 – Germany invades Poland, WWII starts

Alan Turing

- Leads British effort to crack Enigma
- Use cribs ("WETTER" transmitted every day at 6am) to find structure of plugboard settings
- 10,000 people worked at Bletchley Park on breaking Enigma (100,000 for Manhattan Project)

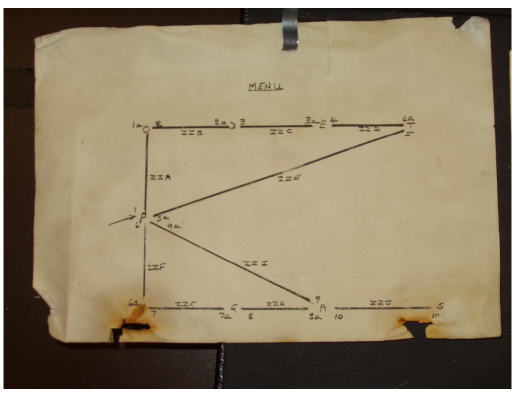
Alan Turing's "Bombe"



Steps through all possible rotor positions (26^3), testing for probable plaintext; couldn't search all plugboard settings (> 10^{12}); take advantage of loops in cribs

"Bombes"

- Idea by Alan Turing
- Name from Rejewski's "Bomba" machine (Polish for bomb)
  - "for lack of a better idea" (Rejewki's paper)
- Design by "Doc" Keen, British Tabulating Machine Co.
- First machine, "Victory": Bletchley Park, March 1940



Enigma Cryptanalysis

- Relied on combination of sheer brilliance, mathematics, espionage, operator errors, and hard work
- Huge impact on WWII
  - Britain knew where German U-boats were
  - Advance notice of bombing raids
  - But...keeping code break secret more important than short-term uses or giving credit: Turing's Enigma report declassified in 1996!

Turing after the War

- Made several major contributions to Computer Science (both before and after)
  - Most important award is named "Turing Award"
- Prosecuted for homosexuality
  - Illegal in Britain
  - Forced hormone treatment
- 1954 – died of cyanide poisoning from eating apple (believed to be suicide)

Next Class: Modern Crypto

- Strong Symmetric Ciphers
  - How they are similar and different
  - How hard to break
- How two people who have never met can communicate securely
  - Public-key Cryptography
- What it means when you see the key symbol on your web browser