Lecture 3: Captain Ridley's Shooting Party

Confronted with the prospect of defeat, the Allied cryptanalysts had worked night and day to penetrate German ciphers. It would appear that fear was the main driving force, and that adversity is one of the foundations of successful codebreaking.

Simon Singh, *The Code Book*

CS588: Security and Privacy
University of Virginia
Computer Science

David Evans
http://www.cs.virginia.edu/~evans

Menu

- Projects
  - Phil Varner
- Vigenère – "Le Chiffre Indéchiffrable"
- Enigma

Projects

- Preliminary Proposals due Oct 1
- Team assignments will be decided by Friday
- Open ended – proposal will lead to an "agreement"
- Different types of projects:
  - Design/Implement
  - Analyze
  - Research Survey
- List of ideas on web
  - Just a starting point, don't limit yourself to ideas on list

Project Evaluation

- Need not be 100% technical: politics, psychology, law, ethics, history, etc.; but shouldn't be 0% technical.
- Design/Implementation projects less focus on quality and organization of writing (but still important)
- All team members get same project grade
  - Unless there are problems: tell me early!

Vigenère

- Invented by Blaise de Vigenère, ~1550
- Considered unbreakable for 300 years
- Broken by Charles Babbage but kept secret to help British in Crimean War
- Attack discovered independently by Friedrich Kasiski, 1863.

Vigenère Encryption

Key is a N-letter string.

$E_K(P) = C$ where

$C_i = (P_i + K_{i \bmod N}) \bmod Z$

(size of alphabet)

$E_{\text{"KEY"}}$("test") = DIQD

C_0 = ('t' + 'K') mod 26 = 'D'

C_1 = ('e' + 'E') mod 26 = 'I'

C_2 = ('s' + 'Y') mod 26 = 'Q'

C_3 = ('t' + 'K') mod 26 = 'D'

Babbage's Attack

- Use repetition to guess key length:
  Sequence XFO appears at 65, 71, 122, 176.
  Spacings = (71 – 65) = 6 = 3 * 2
  (122 – 65) = 57 = 3 * 19
  (176 – 122) = 54 = 3 * 18
  Key is probably 3 letters long.

Key length - Frequency

- Once you know key length, can slice ciphertext and use frequencies:

L_0: DLQLCNSOLSQRNKGBSEVYNDOIOXAXYRSOSGYKY
VZXVOXCDNOOSOCOWDKOOYROEVSRBXENI

Frequencies: O: 12, S: 7, Guess O = e

$C_i = (P_i + K_{i \bmod N}) \bmod Z$

'O' = ('e' + K_0) mod 26

14 = 5 + 9 => K_0 = 'K'

Sometimes, not so lucky...

L_1: LMISQITVYJSSSJAHYECYSXOXGWYGJMRRXEGWRPEJXSI
SLIGHTVSXILWHXYXJPERISWTM

S: 9, X: 7; I: 6 guess S = 'e'

'S' = ('e' + K_1) mod 26

19 = 5 + 14 => K_1 = 'N'

'X' = ('e' + K_1) mod 26

24 = 5 + 19 => K_1 = 'M'

'I' = ('e' + K_1) mod 26

10 = 5 + 5 => K_1 = 'E'

"Recover" one password for ~$35-50

Vigenère Simplification

- Use binary alphabet:
  $C_i = (P_i + K_{i \bmod N}) \bmod 2$
  $C_i = P_i \oplus K_{i \bmod N}$
- Use a key as long as P:
  $C_i = P_i \oplus K_i$
- One-time pad – perfect cipher!

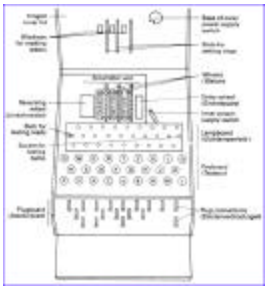
Enigma

Enigma

Enigma machine at NSA Museum

- Invented commercially, 1923
- Adopted by Nazi's
- About 50,000 in use
- Modified throughout WWII, believed to be perfectly secure
- [Kahn67] didn't know it was broken
- Turing's 1940 Treatise on Enigma declassified in 1996.

Enigma Mechanics



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

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 – start orientations of the 3 rotors
  - Ring settings – when next ring advances
- Reflector
  - Fixed symmetric substitution ($A \rightarrow B \Rightarrow B \rightarrow A$)

Key Space

- If you don't know rotors:
  $(26!)^3 \approx 4 * 10^{26}$
- If you don't know reflector:
  $(26 * 25 / 2) * (24 * 23 / 2) * \ldots * (2 * 2) / 13!$
  $\approx 8 * 10^{12}$
- Plugboard with 6 plugs:
  $(26 * 25/2) * \ldots * (16*15 / 2) / 6! \approx 10^{11}$
- Ring settings: $26^2 = 676$
- Message Key: $26^3 = 17576$
- Total: $\approx 6 * 10^{110}$
  (there are 10^{84} atoms in the universe)

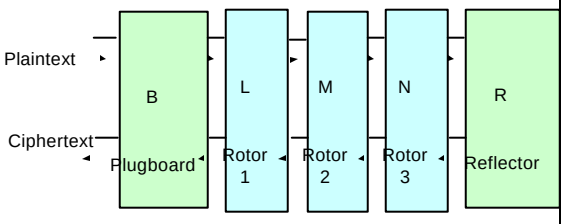
Capturing a Machine

- If you don't know rotors:
  $(26!)^3 \approx 4 * 10^{26}$
- If you don't know reflector:
  $(26 * 25 / 2) * (24 * 23 / 2) * \ldots * (2 * 2) / 13!$
  $\approx 8 * 10^{12}$
- Plugboard with 6 plugs:
  $(26 * 25/2) * \ldots * (16*15 / 2) / 6! \approx 10^{11}$
- Ring settings: $26^2 = 676$
- Message Key: $26^3 = 17576$

Know rotors:
3! = 6
Or 3 from 5 = 5 * 4 *3
= 60

Know reflector: 1

Enigma Schematic



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

Does Decryption Work?

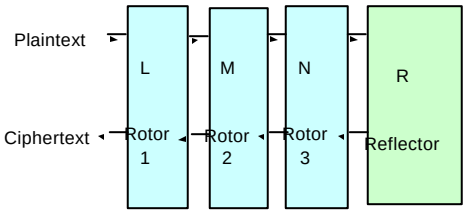
$C = B^{-1}L^{-1}M^{-1}N^{-1}RNMLB(P)$

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

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

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

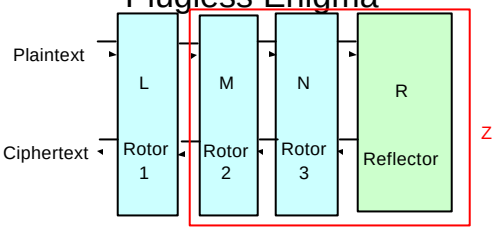
Plugless Enigma



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

Used in Spanish Civil War 1938-39 (by all sides)

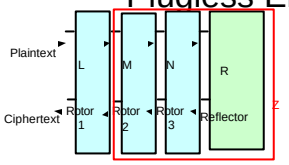
Plugless Enigma



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

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

Plugless Enigma



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

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

$L_1(X) = Z\,L_1(W)$
$L_2(T) = Z\,L_2(E)$
$L_3(S) = Z\,L_3(H)$
$L_4(W) = Z\,L_4(R)$
$L_5(V) = Z\,L_5(M)$
$L_6(U) = Z\,L_6(A)$
$L_7(I) = Z\,L_7(C)$

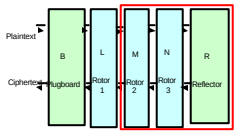
```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
EKMFLGDQVZNTOWYHXUSPAIBRCJ
JEKMFLGDQVZNTOWYHXUSPAIBRC
CJEKMFLGDQVZNTOWYHXUSPAIBR
RCJEKMFLGDQVZNTOWYHXUSPAIB
BRCJEKMFLGDQVZNTOWYHXUSPAI
IBRCJEKMFLGDQVZNTOWYHXUSPA
AIBRCJEKMFLGDQVZNTOWYHXUSP
```

For a given starting rotor setting, solve for Z
1: $R = Z(B)$ 2: $S = Z(F)$ 3: $X = Z(G)$ 4: $P = Z(Y)$
5: $U = Z(V)$ 6: $H = Z(I)$ 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
- Catalog to map Z to rotor settings for R2 and R3

Enter the Plugboard



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

6 plugs: $(26*25)/2 * (24*23)/2 * \ldots * (16*15/2) / 6!$
$= 10^{11}$ times harder (at least)

Ideas for making Batons attack harder without plugboard?

Operation

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

Repeated Message Key

$P = P_1P_2P_3P_1P_2P_3$

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

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

$E_4E_1(C_1) = E_4(P_1) = C_4$
$E_4E_1 = B^{-1}L_1^{-1}M^{-1}N^{-1}RNML_1B\,B^{-1}L_4^{-1}M^{-1}N^{-1}RNML_4B$
$= B^{-1}L_1^{-1}M^{-1}N^{-1}RNML_1L_4^{-1}M^{-1}N^{-1}RNML_4B$

Letter Permutations

Symmetry of Enigma:
if $E_{pos}(x) = y$ we know $E_{pos}(y) = x$
Given message openings

DMQ VBM	$E_1(m_1) = D$	$E_4(m_1) = V$
VON PUY	$E_1(m_2) = V$	$E_4(m_2) = P$
PUC FMQ		

With enough message openings, we can build complete cycles for each position pair:
E_1E_4 = (DVPFKXGZYO) (EIJMUNQLHT) (BC) (RW) (A) (S)
Note: Cycles must come in pairs of equal length
(Examples in Code Book had pairs of unequal length)

Composing Involutions

- E_1 and E_4 are involutions $(x \rightarrow y \Rightarrow y \rightarrow x)$
- Without loss of generality, we can write:
  E_1 contains $(a_1a_2)\ (a_3a_4)\ \ldots\ (a_{2k-1}a_{2k})$
  E_2 contains $(a_2a_3)\ (a_4a_5)\ \ldots\ (a_{2k}a_1)$

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 $x = a_1$

Rejewski's Theorem

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

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

- The product of two involutions consists of pairs 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)

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

$\Rightarrow$ About 20^3 (8000) factorizations to try

(still too many in pre-computer days)

Luckily…

- Operators picked guessable 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, knew plugboard settings (didn't change frequently)

Six equations, two unknowns – is solvable

1939

- Early 1939 – Germany changes scamblers and adds extra plugboard cables, stop double-transmissions
  - Poland unable to cryptanalyze
- July 1939 – Rejewski invites French and British cryptographers
  - It is actually breakable
  - Gives England replica Enigma machine constructed from plans

Bletchley Park

- Alan Turing leads British effort to crack Enigma
- Use cribs ("WETTER" transmitted every day at 6am)
- Still needed to brute force check ~1M keys.
- Built "bombes" to automate testing
- 30,000 people worked at Bletchley Park on breaking Enigma – 100,000 for Manhattan Project

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

Charge

- Next time: Block Ciphers
- Problem Set 1 Due Monday
- Start thinking about projects