Luštění německého šifrovacího stroje Lorenz

podle bakalářské práce Petra Veselého, MFF UK

21. února 2011

- 2000 – zveřejnění dobové zprávy *General Report on Tunny*
- informací nedostatek k odvození konstrukce šifrátoru Lorenz
- **cíl**: odvození pravděpodobného postupu kryptoanalytiků

Program

Program

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

- *Bletchley Park* - kryptoanalytické středisko 80 km SZ od Londýna, založené 1939, existence tajena do 70. let 20. stol.
- červen 1941 - zachycení 1. komunikace pomocí šifrátoru Lorenz (TUNNY)
- další zprávy $\Rightarrow$ šifra Vernamova typu
- 30. 8. 1941 - zachycení 2 téměř stejných zpráv šifrovaných stejným klíčem - rozluštěny plukovníkem Johnem H. Tiltmanem $\Rightarrow$ 3976 znaků pseudonáhodného klíče
- leden 1942 - rekonstrukce šifrátoru Williamem T. Tuttem
- prosinec 1943 - k určování nastavení rotorů vyvinut Colossus - 1. částečně programovatelný počítač na světě
- 8. 5. 1945 - zachycena poslední zpráva
- celková délka zpráv 63 431 000 znaků
- úspěchy díky luštění Lorenze viz. Crypto-World 2008

Program

- Lorenz versus Enigma:
  - Enigma - díky přenosnosti výbava bojových útvarů nejnižší úrovně
  - Lorenz - linky mezi nejvyšším velitelstvím pozemní armády v Berlíně a hlavními stany armádních skupin v Evropě a S Africe
- červen 1941 - říjen 1942 - zkušební provoz na lince Berlín, Soluň, Athény
- říjen 1942 - ostré vysílání Berlín, Soluň a Královec, J Rusko
- 1944 - síť 26 linek

Program

- Lorenz versus Enigma:
  - Lorenz SZ40 (SZ42A, SZ42B) “Schlüsselzusatzgerät”, tj. přídavný modul k bezdrátovému dálnopisu (šifrování on-line)
  - Enigma - předběžné šifrování zpráv (off-line) a následné odeslání běžným komunikačním kanálem
- dálnopis = telekomunikační zařízení ($\sim$psací stroj), elektronický přenos zpráv po lince nebo bezdrátově a tisk
- Baudotův dálnopisný kód (str. 9) - 5-bitový, konkrétní signál ($\times$, $\cdot$) dle komunikační linky

Program

- skládá se z 12 rotorů (str. 9)
  1. $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_3$, $\mathcal{K}_4$, $\mathcal{K}_5$ délek 41, 31, 29, 26, 23,
  2. $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$, $\mathcal{S}_5$ délek 43, 47, 51, 53, 59,
  3. $\mathcal{M}_1$, $\mathcal{M}_2$ délek 61, 37
- kolíčky na rotorech ve 2 možných polohách: 0 a 1

Program

- generuje pseudonáhodný klíč
- v každém kroku 5-bitový znak klíče: i-tý impuls = součet aktivních kolíčků $\mathcal{K}_i + \mathcal{S}_i \mod 2$
- pravidla pohybu
  1. $\mathcal{K}_i$ se točí vždy
  2. $\mathcal{S}_i$ se točí všechna $\Leftrightarrow$ aktivní kolíček $\mathcal{M}_2$ před případným otočením v poloze 1
  3. $\mathcal{M}_2$ se točí $\Leftrightarrow$ aktivní kolíček $\mathcal{M}_1$ před otočením v poloze 1
  4. $\mathcal{M}_1$ se točí vždy
- společné otáčení S_i je hlavní slabinou Lorenze!!!

Program

- obtížnost:
  1. každý měsíc - změna vzorků kol $\mathcal{K}_i$
  2. každé 3 měsíce - změna vzorků kol $\mathcal{S}_i$
  3. každý den - změna vzorků kol $\mathcal{M}_1, \mathcal{M}_2$
- záchytný bod: 12 písmenný indikátor v hlavičce depeše
  1. stejné indikátory $\Rightarrow$ stejný klíč
  2. nedůslednost operátorů (malá variace indikátorů)
  3. pedantství (zašifrované slovo Spruchnummer na začátku zpráv)

- aditivní šifra Vernamova typu ($\mathcal{C} = \mathcal{M} + K$, a tedy $\mathcal{M} = \mathcal{C} + K$)
- slabina: zprávy šifrované stejným klíčem $\mathcal{M} - \mathcal{M}' = \mathcal{C} - \mathcal{C}'$ (rozluštění jedné zprávy vede k rozluštění další šifrované stejným klíčem)
- 30. 8. 1941 - zachyceny 2 zprávy se stejným indikátorem HQIBPEXEZMUG, kratší o délce 3976 znaků
- až na zkratky, interpunkci, překlepy tatéž zpráva (identická zpráva $\neq$ kompromitace!)
- rozluštěna během dvou měsíců plukovníkem John H. Tiltmanem $\Rightarrow$ získáno **3976 znaků pseudonáhodného klíče**

Program

- pozorování: indikátor ABCDEFGHIJKL a BBCDEFGHIJKL - příslušné klíče v Baudotově kódu stejné až na 1. impuls

$$K = \begin{pmatrix}0\\1\\1\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}\ldots \text{ a } K' = \begin{pmatrix}1\\1\\1\\1\\0\end{pmatrix}\begin{pmatrix}1\\0\\1\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\1\\0\\1\end{pmatrix}\ldots$$

každý znak klíče = 5 impulsů, tj. místo K zkoumáme binární posloupnosti K_1, K_2, K_3, K_4, K_5

K :	C	W	V	S	4	...
K_1 :	0	1	0	1	0	...
K_2 :	1	1	1	0	1	...
K_3 :	1	0	1	1	0	...
K_4 :	1	0	1	0	0	...
K_5 :	0	1	1	0	0	...

Hledání periody $\mathcal{K}_1$

- v K_1 hledáme opakující se posloupnosti (Kasiského test)
- nejdelší má 26 znaků (str. 19), shodné úseky vzdáleny o násobky 41
- hypotéza: $K_1 = \mathcal{K}_1 + \mathcal{S}'_1 \mod 2$, kde $\mathcal{K}_1$ periodická (s periodou 41) a $\mathcal{S}'_1$ má dlouhé opakující se úseky
- analogie Vigenèrovy šifry nad abecedou $\{0, 1\}$, kde šifrovým textem je K_1, klíčem je prvních 41 znaků posloupnosti $\mathcal{K}_1$ a otevřeným textem je $\mathcal{S}'_1$

Ověření periody $\mathcal{K}_1$

- *index koincidence*: dán text $T = t_1 \dots t_m$

  $$IC(T) = \mathrm{p}\left\{t_i = t_j \;\middle|\; i, j \text{ libovolně zvolené, } 1 \leq i < j \leq m\right\}$$

- určení periody Vigenèrovy šifry
  1. šifrový text do tabulek o různém počtu sloupců
  2. IC textů ve sloupcích
  3. průměrný IC tabulek
  4. počet sloupců tabulky s maximálním IC = perioda
  5. IC sloupců = IC otevřeného textu

Index koincidence $\mathcal{S}'_1$

- IC v K_1 pro bigramy, protože 0 a 1 rozděleny rovnoměrně
- tabulka o l sloupcích a IC dvojic sloupců $1;2,\ 2;3,\ \ldots\ l-1;l$
- pro $2 \leq l \leq 99$ určíme průměrný IC
- maximální pro $l = 41$ a $l = 82 \Rightarrow$ perioda $\mathcal{K}_1$ je 41
- analogicky periody $\mathcal{K}_i$

$$\begin{aligned} i = 1: &\quad 41 \\ i = 2: &\quad 31 \\ i = 3: &\quad 29 \\ i = 4: &\quad 26 \\ i = 5: &\quad 23 \end{aligned}$$

- odhalení nerovnoměrného rozložení bigramů v $\mathcal{S}'_i$!

Hledání $\mathcal{K}_5$

- $\mathcal{K}_5 = k_1 k_2 k_3 \ldots k_{23}$ má nejkratší periodu 23
- předpoklad BÚNO: $\mathcal{S}'_5$ převládají bigramy $(0,0)$ a $(1,1)$ $\Rightarrow$ převládají-li v dvojici sloupců $(i, i+1)$ (v K_5 zapsaném do tabulky o 23 sloupcích)
  1. bigramy $(0,0)$ a $(1,1)$, pak $(k_i, k_{i+1}) \in \{(0,0),(1,1)\}$
  2. bigramy $(0,1)$ a $(1,0)$, pak $(k_i, k_{i+1}) \in \{(0,1),(1,0)\}$
- pro 1. a 2. dvojici sloupců převaha $(0,1)$ a $(1,0)$ (str. 26) $\Rightarrow$ platí

$$(k_1, k_2) + (k_2, k_3) \in \{(0,0),(1,1)\},$$

  a tedy $k_3 = k_1$
- pro 1. a 3. dvojici sloupců

$$(k_1, k_2) + (k_3, k_4) \in \{(0,0),(1,1)\},$$

  a tedy $k_2 = k_4$ atd.
- pro $k_1 := 0$ je $\mathcal{K}_5 = 0101111000101110001101$

Analýza $\mathcal{S}'_i$

$\mathcal{S}'_1$:	0	0	1	1	0	0	0	0	1	1	1	1
$\mathcal{S}'_2$:	1	1	1	1	0	0	0	0	1	0	0	0
$\mathcal{S}'_3$:	1	1	0	0	1	1	1	1	0	0	0	0
$\mathcal{S}'_4$:	0	0	1	1	1	1	1	0	0	1	1	1
$\mathcal{S}'_5$:	1	1	0	0	0	0	1	0	0	0	0	0
$\mathcal{S}'$:	P	P	J	J	N	N	M	9	A	D	D	D

- časté opakování písmen v $\mathcal{S}' \Rightarrow S_i$ rotory, které často stojí
- 00111000010 obsahuje běhy $00, 111, 0000, 1, 0$
- S'_i vzniká z S_i prodloužením některých běhů S_i
- 2 úlohy
  1. určení počtu běhů ve vzorcích kol S_i
  2. zjištění délky každého běhu

Počet běhů v $\mathcal{S}_i$

- hlavní myšlenka: $\mathcal{S}'_i$ má běh délky 1 (010 nebo 101 je v $\mathcal{S}'_i$), pak $\mathcal{S}_i$ má běh délky 1
- z četnosti výskytů bigramů v $\mathcal{S}'_i \Rightarrow$ takových běhů málo
- v $\mathcal{S}'_1$ jsou posloupnosti tvaru $0\underline{10\ldots0}10$
- počet běhů v nich je násobek 18 $\Rightarrow$ zřejmě je 18 i počet běhů v kole $\mathcal{S}_1$

Délky běhů v $\mathcal{S}_i$

- výpis posloupností tvaru $0\underline{10\ldots10}$ do řádků pod sebe
  101100000111110001111100111111000011000011111100011000011000
  101100001111111000011111001111110000011000000110001111100011100
  101111000011111000111000111111000011110000001110011110000111100
  ...
- výpis délek jejich běhů do řádků pod sebe
  1125535264346334 23
  1124745265262353 32
  1144533364463234 42
  ...
- v každém sloupci minimum = délka příslušného běhu v $\mathcal{S}_1$
- $\mathcal{S}_1$ tak určena jednoznačně až na počáteční počet nul

$$\mathcal{S}_1 = 011000110011100011001011000111001110011110$$

nebo

$$\mathcal{S}_1 = 001100011001110001100101100011100111001111 0$$

Odvození řídicích posloupností

- definujeme řídicí posloupnosti $\mathcal{R}_i = \{r_j^{(i)}\}_{j=1}^{3975}$

$$r_j^{(i)} = 1 \quad \text{na konci } j\text{-tého kroku se } \mathcal{S}_i \text{ točí}$$
$$r_j^{(i)} = 0 \quad \text{na konci } j\text{-tého kroku } \mathcal{S}_i \text{ stojí}$$

- jen částečná rekonstrukce $\mathcal{R}_i$ z $\mathcal{S}_i$ a $\mathcal{S}'_i$:

$$\begin{aligned}
\mathcal{S}_1 &= (0)01100011001110001100101100011100111001111 0 \\
\mathcal{S}'_1 &= 00110000111100111111000000110001110011000 0 \ldots \\
\mathcal{R}_1 &= ?110111\ldots \\
\mathcal{R}_1 &= ?111011\ldots \\
\mathcal{R}_1 &= ?111101\ldots
\end{aligned}$$

- posloupnosti $\mathcal{R}_i$ (str. 39)

Odvození řídicích posloupností

- 12 písmenný indikátor $\Rightarrow$ 12 rotorů $(\mathcal{K}_i, \mathcal{S}_i, \mathcal{M}_1, \mathcal{M}_2) \Rightarrow$ společné otáčení kol $\mathcal{S}_i$
- předpoklad: $\mathcal{R} = \mathcal{R}_i \Rightarrow$ žádný spor, téměř celá $\mathcal{R}$ (str. 40)
- Kasiského test $\Rightarrow$ $\mathcal{R}$ periodická s periodou $2257 = 37 \cdot 61 \Rightarrow$ 61 a 37 kandidáti na délky kol $\mathcal{M}_1$ a $\mathcal{M}_2$ řídících otáčení $\mathcal{S}_i$
  1. 1. nápad: $\mathcal{M}_1$ a $\mathcal{M}_2$ se točí společně $\Rightarrow$ soustava pro $37 + 61$ neznámých $\mathcal{R} = \mathcal{M}_1 + \mathcal{M}_2$, ale nemá řešení
  2. 2. nápad: kolo $\mathcal{M}_2$ se řídí pohybem kola $\mathcal{M}_1$, konkrétně: $\mathcal{M}_1$ se točí vždy a $\mathcal{M}_2$ jen tehdy, je-li na $\mathcal{M}_1$ aktivní kolíček v poloze 1, $\mathcal{R}$ vznik prodloužením některých běhů v $\mathcal{M}_2$

Odvození řídicích posloupností

- $\mathcal{R}$ obsahuje dost osamocených nul, ale málo osamocených jedniček
- $\mathcal{R}$ obsahuje 10100101 (hned 2 osamocené jedničky), z řídkosti výskytu hypotéza: vzorek kola $\mathcal{M}_2$ obsahuje 2 osamocené jedničky jednou
- počet běhů mezi po sobě jdoucími výskyty 2 osamocených jedniček $=$ násobek 24, $\mathcal{R}$ do řádků po 24 bězích a odvodíme vzorek (očekávaných 37 znaků)

$$1010111011101110101101110111011010110$$

- $\mathcal{M}_1$ periodická s periodou 61
- $\mathcal{M}_1$ plyne ze znalosti $\mathcal{M}_2$ a $\mathcal{R}$
- naopak z $\mathcal{M}_1$ dopočteme $\mathcal{R}$, a tedy i počáteční nastavení rotorů $\mathcal{S}_i$

Kolo	Vel.	Vzorek
$\mathcal{K}_1$	41	01100111000011100111000010011011000110110
$\mathcal{K}_2$	31	00011001100010111101110000011011
$\mathcal{K}_3$	29	01111011000111000110001100100
$\mathcal{K}_4$	26	01100010110011100110101001
$\mathcal{K}_5$	23	01011110001011100001101
$\mathcal{S}_1$	43	01100011001110001100101100011100111001111100
$\mathcal{S}_2$	47	11100010001101001110011100110001111000111001100
$\mathcal{S}_3$	51	100111001100101110011000110001110001000111100010011
$\mathcal{S}_4$	53	10000110001100110001101000110011100110110001100111011
$\mathcal{S}_5$	59	01110111000110100111011000110011000110011100001101100011000
$\mathcal{M}_1$	61	1011101101101101101110101011101110101101101010111101010111011
$\mathcal{M}_2$	37	1101110111011010110101011101110111010