Podstawy kryptografii

Andrzej M. Borzyszkowski

Instytut Informatyki
Uniwersytet Gdański

sem. letni 2025/2026

inf.ug.edu.pl/~amb/

Kryptografia klasyczna c.d.

Szyfr monoalfabetyczny

- szyfr monoalfabetyczny (kod), np.

  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
  q w e r t y u i o p a s d f g h j k l z x c v b n m\\

  – kryptoanaliza: analiza częstotliwości wystąpień liter
  – ale można mieć kilka odpowiedników dla każdej litery (homofonia)
  – np. $E : \{A..Z\} \to \{00..99\}$, $|E(e)| = 12$, $|E(z)| = 1$
  – można/trzeba też rozważać częstotliwości par liter, niektóre nie występują praktycznie wcale, inne b. często

Szyfr podstawieniowy, książka kodowa

- dwie książki z parami tekst jawny – tekst zaszyfrowany

Februar	13605
fest	13722
finanzielle	13850
folgender	13918
Frieden	17142

  – w jednej kolejność tekstu jawnego, w drugiej zaszyfrowanego
  – kluczem jest para książek! niesłychanie trudno o wymianę
- wersja: dodatkowym kluczem jest przesunięcie
  – $E(M) = \text{Książka}(M) + \text{przesunięcie} \mod 100.000$
  – $C(C) = \text{Książka}^{-1}(M - \text{przesunięcie} \mod 100.000)$

Szyfr Vigenere'a

klucz: wektor liczb nieznanej długości np. $(k_1, k_2, \ldots, k_9)$

- szyfrowanie: podobnie do Cezara: $y_1 = x_1 + k_1$, $y_2 = x_2 + k_2$, ..., $y_9 = x_9 + k_9$, odszyfrowywanie analogicznie
- kryptoanaliza: przeszukiwanie wyczerpujące jest nierealne, kluczy jest $26 \cdot 26 \cdot \ldots$, dla $n = 9$ jest ich $5 \cdot 10^{12}$
- para tekst jawny+zaszyfrowany, długości klucza, definiuje klucz
- gdy znany jest tylko szyfrogram oraz długość klucza, to można/należy przeprowadzić analizę częstotliwości dla fragmentów szyfrogramu, dla zestawów $\{y_1, y_{10}, y_{19}, \ldots\}$ itd.
- analiza częstotliwości oznacza przybliżenie wektora częstotliwości wystąpień liter w szyfrogramie z częstotliwościami języka naturalnego

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
8	2	3	4	12	2	2	6	7	0	1	4	2	7	7	2	0	6	6	9	3	1	2	0	2	0

Szyfr Vigenere'a, c.d.

znalezienie długości klucza:

- niech $A_0 = (p_0, p_1, p_2, \ldots, p_{25})$ oznacza prawdopodobieństwa występowania liter w tekście
- niech $A_i = (p_i, p_{i+1}, \ldots, p_{25}, p_0, \ldots, p_{i-1})$ oznacza ten sam wektor z przesuniętymi wielkościami
- testujemy prawdopodobieństwo powtórzenie się litery w szyfrogramie oraz szyfrogramie przesuniętym o n miejsc
- jeśli $n =$ długość klucza, to litery były szyfrowane tym samym przesunięciem, prawd $= p_0 \cdot p_0 + p_1 \cdot p_1 + \ldots + p_{25} \cdot p_{25} = A_0 * A_0$
- jeśli $n <>$ długość klucza, to koincydencje przypadają na różne przesunięcia, prawdopodobieństwo koincydencji jest uśrednione po różnych iloczynach $A_0 * A_1$, $A_0 * A_2$, itd
- iloczyn $A_0 * A_0$ jest znacząco większy od innych (w jęz. ang. $A_0 * A_0 \approx 0.066$, inne iloczyny są równe ok. $\frac{1}{26}$

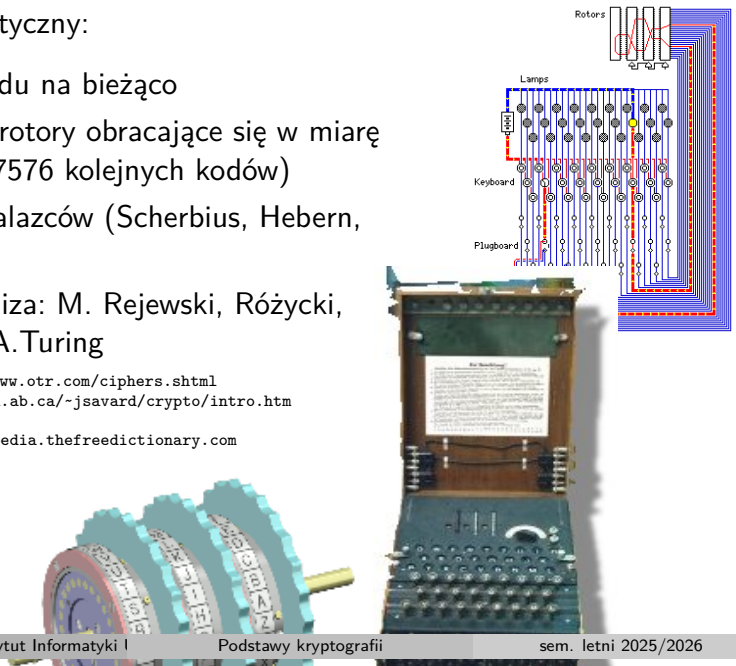
Enigma]

szyfr polialfabetyczny:

- zmiana kodu na bieżąco
- Enigma – rotory obracające się w miarę pisania (17576 kolejnych kodów)
- wielu wynalazców (Scherbius, Hebern, Koch)
- kryptoanaliza: M. Rejewski, Różycki, Zygalski, A.Turing

grafika: http://www.otr.com/ciphers.shtml
http://home.ecn.ab.ca/~jsavard/crypto/intro.htm

http://encyclopedia.thefreedictionary.com



Enigma, kryptoanaliza

- Protokół: wybrać klucz sesyjny (3 litery), powtórzyć go dwukrotnie i zaszyfrować kluczem dziennym
  – dane: setki szyfrogramów o powyższym początku
  – klucz dzienny wyznacza permutacje $P_1, P_2, \ldots, P_6$ (i pozostałych 17570), złożenia $P_4 \circ P_1^{-1}$ i pozostałe dwa są podane implicite w danych
  – np. $ABCABC \rightarrow ENIGMA$, więc $P_4 \circ P_1^{-1}(E) = G$ itd.
  – dodatkowe podstawienie λ unieważnia dokładną znajomość permutacji $\lambda \circ P_4 \circ P_1^{-1} \circ \lambda^{-1}$
  – ale „kształt" (rozkład na cykle) permutacji jest ustalony
  – opracowano enumeratywną listę rozkładów dla 100.000 kluczy

Kryptografia nowoczesna

Zasady nowoczesnej kryptografii

- Precyzyjne definicje
- Jawne założenia systemu
- Dowody poprawności
- (Kontr)przykład: definicja wymagań dla bezpiecznego systemu kryptograficznego
  – nie da się odtworzyć klucza kryptograficznego .....
  – nie da się odtworzyć tekstu jawnego .....
  – nie da się odtworzyć żadnego fragmentu tekstu jawnego
  – nie da się uzyskać żadnej sensownej informacji .....
  – nie da się obliczyć żadnej funkcji zależącej od tekstu jawnego
- dopiero ostatnia definicja w miarę precyzyjnie opisuje wymogi dla systemu kryptograficznego

Założenia/dowody?

- Model matematyczny a świat fizyczny
  – przykład: złamanie szyfru implementowanego na karcie kryptograficznej w oparciu o analizę zużycia energii
- Jawne założenia + analiza spełnienia założeń
- Różnica pomiędzy oprogramowaniem ogólnego użytku a kryptografią
  – kto zauważa błędy systemu/implementacji?
  – kto celowo szuka błędów w systemie/implementacji?

Atak z wybranym tekstem jawnym

- Bezpieczeństwo przeciwko atakowi z kryptogramem
  – Ewa nie jest w stanie zgadnąć, która wiadomość jest zaszyfrowana na podstawie znajomości kryptogramu
- Bezpieczeństwo szyfrowania wielokrotnego
  – Ewa nie potrafi odgadnąć, który zestaw wiadomości jest szyfrowany
  – tw.: szyfrowanie musi być niedeterministyczne
- Bezpieczeństwo przeciwko atakowi z wybranym tekstem jawnym – CPA
  – założenie: Ewa ma dostęp do maszyny szyfrującej (wyrocznia – daje kryptogram dowolnej wiadomości)
  – Ewa nie potrafi odgadnąć, która wiadomość jest zaszyfrowana

Własności ataku z wybranym testem jawnym

- Twierdzenie: szyfr bezpieczny przeciwko atakowi z wybranym tekstem jawnym musi być niedeterministyczny
  – dw.: Ewa żąda od wyroczni zaszyfrowania obu wiadomości i sprawdza, która została podana jej jako wyzwanie
- Twierdzenie: atak z wybranym tekstem jawnym jest łatwiejszy niż z zestawem tekstów jawnych wybranych z góry
- Tryb asynchroniczny vs. synchronizacja
  – szyfr może odwoływać się do stanu
  – zależność od stanu może zastąpić niedeterminizm, każde szyfrowanie tej samej wiadomości da inny wynik, bo jest inny stan
  – w zasadzie rozpatrujemy szyfry nie odwołujące się do pojęcia stanu

Własności ataku z wybranym tekstem jawnym c.d.

- Klasyczne szyfry są prawie zawsze nieodporne na atak z tekstem jawnym
  – szyfr Cezara, Vigenere'a, Hilla
  – wielokrotne przykłady łamania szyfrów w praktyce (atak na Midway, „żądanie" zaszyfrowania słowa 'Midway')
- Odporność na atak z wybranym tekstem jawnym jest ważna
  – serwery odpowiadają na żądania użytkowników
  – być może są to ataki
- Tw.: odporność na atak z wybranym tekstem jawnym = odporność CPA przy wielokrotnym szyfrowaniu