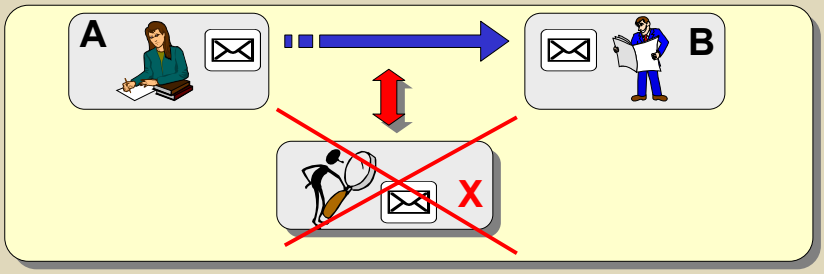
Elementy kryptografii

Szyfrowanie danych przy użyciu kluczy symetrycznych

Szyfrowanie danych = ukrycie treści



Jak utajniać dane?

Albert: „Jak wezmę to pomarańczowe, to znaczy ‘padnij’”

[Seksmisja]

Wady - za dużo współdzielonej informacji, brak reguły

Szyfrowanie danych

“doohd ldfxd hwx”

[Juliusz Cezar]

Szyfr Cezara

‘Nowoczesny’ szyfr - precyzyjny algorytm kodowania + tajna informacja potrzebna do deszyfracji

Zasada kodowania

$c_i = (m_i + k) \bmod n$

k - klucz

c_i- litera kodu, m_i - litera wiadomości, n - liczba znaków alfabetu

x	y	z	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	w

k = 3

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	w	x	y	z

gladiator → j o d g l dx r u

Szyfr Cezara

doohd ldfxd hwx —($k=3$, A→ D, B→ E ...)→ **allea iacta est**

[Juliusz Cezar]

Szyfr Cezara - szyfr podstawieniowy

Jak złamać szyfr Cezara ?
(czyli odczytać zaszyfrowaną wiadomość)

Bezpieczeństwa szyfru Cezara chroni przede wszystkim algorytm szyfrowania

Znając algorytm łatwo odnaleźć k dające sensowną treść wiadomości

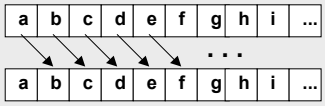
?: nsktxrfzdpf

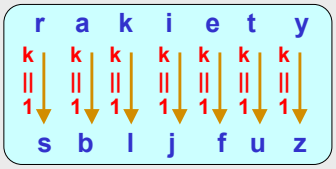
Ewolucja szyfrów podstawieniowych

Zmiana zasady kodowania liter - zmienna wartość przesunięcia

Stała wartość przesunięcia (szyfr Cezara)

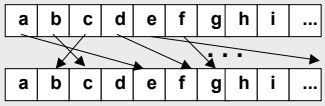
k = 1

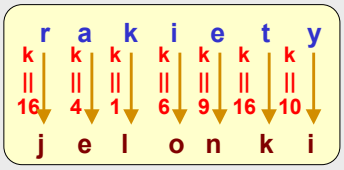




Zmienna wartość przesunięcia

k = var



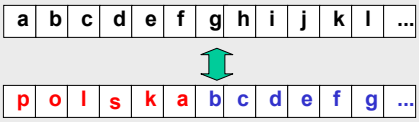


Ewolucja szyfrów podstawieniowych

Jak zapamiętać kody przypisane kolejnym znakom ?

Użyć sekretnego słowa - klucza, np.:

'polska'



Zwiększenie długości klucza - większe zróżnicowanie zasad zastępowania liter

Jak złamać szyfr podstawieniowy z kluczem ?

➡ **Metoda prób i błędów - analiza wszystkich możliwych kluczy**

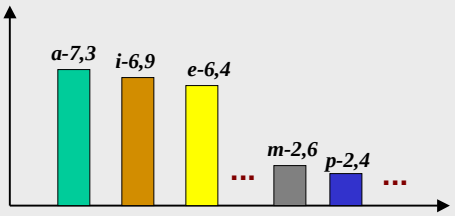
3 literowe = 3! = 6 ; 4 literowe = 4! = 24 ; 5 literowe = 5! = 120 ;

... 14 literowe = 14! = 87 178 291 200 ➡ dla długich kluczy - zadanie nierealne

Szyfry podstawieniowe, c.d.

Bezpieczna metoda szyfrowania ?

Charakterystyczna cecha każdego języka - częstości występowania znaków i grup znaków



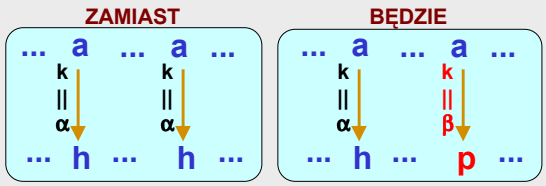
Analiza częstości występowania znaków pozwala na łatwe łamanie szyfrów, w których relacja znak tekstu jawnego - znak kodu jest jednoznacznie określona

Szyfry podstawieniowe monoalfabetyczne nie zapewniają skutecznej ochrony danych

Szyfry wieloalfabetyczne

Cel: ukryć częstość występowania znaków

Zasada kodowania nie jest stała



Zasada szyfrowania

Klucz: N a j l e p s z e k a s z t a n y są na placu Pigalle

Przesunięcie (k): 14 10 5 ..... 0 12

Tekst jawny: Przejmuję władzę jutro po kolacji → Frj...

P → f (k=14) r → r (k=0) z → j (k=10) .....

Szyfry wieloalfabetyczne

Właściwości metody

Metoda maskuje częstości występowania liter i zapewnia teoretyczną nieprzełamywalność szyfru jeżeli

- **klucz jest nie krótszy od wiadomości**
- **kluczem jest ciąg liczb losowych**

W przeciwnym razie klucz staje się 'okresowy', przy czym długość okresu szyfru można oszacować (metoda Kasiskiego)

a wtedy, im krótszy klucz, tym łatwiej złamać szyfr

Szyfry wieloalfabetyczne

Wniosek: trzeba wydłużać długość klucza

Używanie długiego klucza wymaga wsparcia maszyny

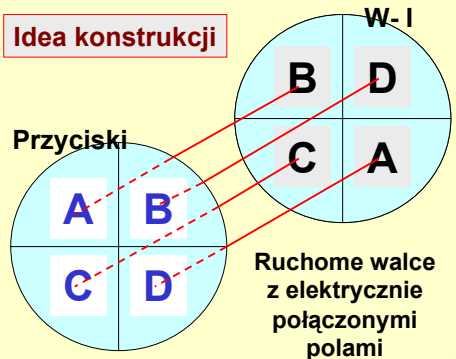
ENIGMA

Pierwsza faza 'wojny' z ENIGMĄ
1933-39 - złamanie szyfru
M.Rejewski, J.Różycki, H.Zygalski

1934 - zapowiedź 'nocy długich noży'
...
1938 - aneksja Czech

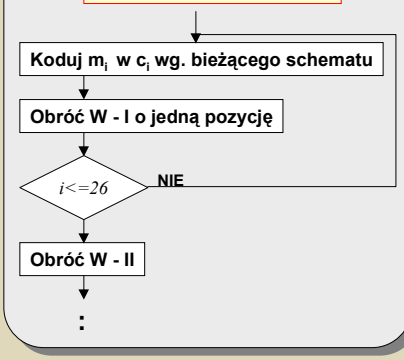
ENIGMA

Idea konstrukcji



Ruchome walce z elektrycznie połączonymi polami

Algorytm działania



Przykład **Tekst jawny - BABA** **Kryptogram - ABBA**

1 B → A 2 obrót 3 A → B 5 B → B 7 A → A

ENIGMA

Podstawowe właściwości

Okres powtarzania klucza - 26^n gdzie n - liczba walców

→ Enigma ‘trójtarczowa’ (do wybuchu wojny cała armia, po - wojska lądowe)

‘Długość’ klucza: $26^3 = 17576$

→ Enigma ‘pięciotarczowa’ (od wybuchu wojny marynarka i lotnictwo)

‘Długość’ klucza: $26^5 = 11\,881\,376$

ENIGMA miała pewną wadę, która wydatnie pomogła w złamaniu szyfru

Współczesna kryptografia

Kryptografia ery komputera

Łamanie szyfrów - podstawowy czynnik stymulujący powstanie komputerów

1940 - A. Turing, 1943 - K. Zeusse (Niemcy) maszyny do "łamania" szyfrów
ENIAC (1943) - podstawowe zadanie - krytpografia

Założenie 1 współczesnej kryptografii

Algorytm szyfrowania jest niemożliwy do utajnienia

Rodzaje "ataków" kryptograficznych

1 Atak z tekstem zaszyfrowanym

Dane: Algorytm, Szyfrogram (C)
Cel: Wiadomość (M), Klucz (K)

2 Atak z tekstem jawnym / z tekstem spreparowanym

Dane: Algorytm, Szyfrogram (C), Wiadomość (M)
Cel: Klucz (K)

Założenie 2 współczesnej kryptografii

Najbardziej prawdopodobny atak ze spreparowanym tekstem jawnym

Bezpieczeństwo szyfru zależy całkowicie od klucza
Odgadnięcie klucza musi być nierealne

Ilościowe charakterystyki szyfrów

Entropia

Miara przypadkowości symboli tekstu

$$H = \sum_i p(x_i) \log \frac{1}{p(x_i)} = -\sum_i p(x_i) \log p(x_i)$$

$p(x_i)$: **prawdopodobieństwo pojawienia się symbolu** x_i

Przykład *systemy sieciowe*

Symbole (x_i)

x_0 - 's', x_1 - 'y', x_2 - 't', x_3 - 'e', x_4 - 'm',
x_5 - 'y', x_6 - 'i', x_7 - 'c', x_8 - 'o', x_9 - 'w'

Prawdopodobieństwa ($p(x_i)$)

$$p_0 = \frac{3}{15} \quad p_1 = \frac{2}{15} \quad p_2 = \frac{1}{15} \quad p_3 = \frac{3}{15} \quad p_4 = \frac{1}{15}$$
$$p_5 = \frac{2}{15} \quad p_6 = \frac{2}{15} \quad p_7 = \frac{1}{15} \quad p_8 = \frac{1}{15} \quad p_9 = \frac{1}{15}$$

Entropia

$$H = -(2 \cdot \frac{3}{15} \cdot \log \frac{15}{3} + 3 \cdot \frac{2}{15} \cdot \log \frac{15}{2} + 5 \cdot \frac{1}{15} \cdot \log \frac{15}{1}) = -1.021$$

Ilościowe charakterystyki szyfrów

Porównanie entropii wiadomości szyfrowanej przy użyciu szyfrów mono- i wielo-alfabetycznych

systemy sieciowe → **szyfr monoalfabetyczny** $H_1 = ?$

systemy sieciowe → **szyfr wieloalfabetyczny** $H_2 = ?$

systemy sieciowe

Szyfr monoalfabetyczny - klucz 'polska' (bez polskich czcionek)

$$p_0 = \frac{3}{15} \quad p_1 = \frac{2}{15} \quad p_2 = \frac{1}{15} \quad p_3 = \frac{3}{15} \quad p_4 = \frac{1}{15}$$
$$p_5 = \frac{2}{15} \quad p_6 = \frac{2}{15} \quad p_7 = \frac{1}{15} \quad p_8 = \frac{1}{15} \quad p_9 = \frac{1}{15}$$

ryrtkhy rdkldjwk ⇒ **Entropia** $H_1 = -1.021$

Ilościowe charakterystyki szyfrów

systemy sieciowe

Szyfr polialfabetyczny - klucz ‘polska’ (bez polskich czcionek)

$$p_0=\frac{1}{15}\quad p_1=\frac{1}{15}\quad p_2=\frac{2}{15}\quad p_3=\frac{1}{15}\quad p_4=\frac{2}{15}$$
$$p_5=\frac{2}{15}\quad p_6=\frac{1}{15}\quad p_7=\frac{1}{15}\quad p_8=\frac{1}{15}\quad p_9=\frac{1}{15}$$
$$p_{10}=\frac{1}{15}\quad p_{11}=\frac{1}{15}$$

jnflpyo iwxniglr ⇒ Entropia $H_2 = -0.98$

$H_2 > H_1$

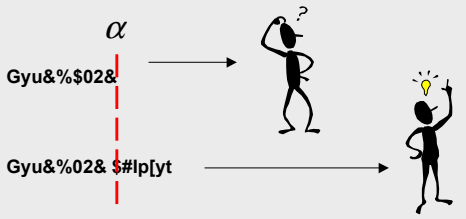
Im większa jest entropia wiadomości, tym lepszy jest szyfr

Algorytm szyfrowania powinien zapewnić maksymalną entropię przesyłanego szyfrogramu

Bezpieczeństwo systemów informatycznych

Ilościowe charakterystyki szyfrów

Długość krytyczna szyfru α



Poufność doskonała: $\alpha \rightarrow \infty$ (szyfr z kluczem jednokrotnym)

Elementy kryptografii Krzysztof Ślot @ 2002

Szyfry przestawieniowe (anagramy)

Idea szyfrowania - zamiana miejscami znaków tekstu

Rodzaje szyfrów przestawieniowych

- **Szyfry bez kluczy**
  Tajność wiadomości zapewnia wyłącznie algorytm szyfrowania
  Przykłady - szyfr 'parkanowy'

- **Szyfry z kluczem**
  Tajność wiadomości jest chroniona przez sekretną informację
  Przykłady - Playfair (I wojna światowa)

Szyfry przestawieniowe

Szyfr 'parkanowy'

Zasada tworzenia kryptogramu

Tekst jawny: Główną gonitwę wygra Tornado

- **Co drugi znak przepisujemy do drugiej linii,**

Góngntęyrtrao
łwąoiwwgaond

- **Łączymy utworzone w ten sposób ciągi**

Kryptogram: Góngntęyrtraołwąoiwwgaond

Szyfry przestawieniowe

Publiczne deponowanie sekretu

Problem - jak opublikować sekret, a jednocześnie go nie wyjawić?

Metoda - anagramy grupujące alfabetycznie litery tekstu

Tajemnica Jutro spadnie deszcz	⇒	Kryptogram acddeeijnoprsstuzz

Właściwości szyfrów przestawieniowych

- ⇒ analiza częstości nie pomaga rozszyfrować wiadomości
- ⇒ są tak samo trudne do złamania jak szyfry podstawieniowe wieloalfabetyczne

Podsumowanie

- ⇒ **Rodzaje technik szyfrowania**
  - Podstawienia
  - Przestawienia
- ⇒ **Założenia współczesnej kryptografii**
  - Jawność algorytmu szyfrowania
  - Jawność pary: tekst - szyfrogram
- ⇒ **Najważniejszy element szyfru**
  - Klucz

Klucz do szyfrowania i do deszyfracji taki sam ⇒ **Klucz symetryczny**

LPMJFD