

Kryptografia

Wprowadzenie do kryptografii

dr Robert Borowiec

Politechnika Wrocławska

Instytut Telekomunikacji i Akustyki

pokój 908, C-5

tel. 3203083

e-mail: robert.borowiec@ita.pwr.wroc.pl

www: istwww.ita.pwr.wroc.pl/~RB/

Wykład IV

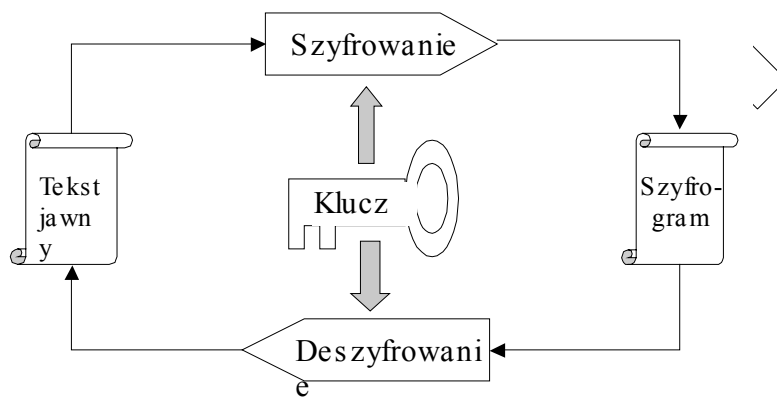
2-godziny

Pojęcia podstawowe

- ⇒ *Kryptografia* jest gałęzią wiedzy i badań zajmującą się utajnionym zapisywaniem informacji;
- ⇒ *Kryptoanaliza* jest dziedziną wiedzy i badań zajmującą się metodami przełamывania szyfrów.
- ⇒ *kryptoanaliza + kryptografia = kryptologia*;
- ⇒ *Szyfrowaniem* nazywamy metodę utajnionego zapisywania *tekstu jawnego* w postaci *tekstu zaszyfrowanego (kryptogramu, szyfrogramu)*;
- ⇒ *Deszyfrowaniem* nazywamy proces przekształcania *szyfrogramu* w *tekst jawny*;

Pojęcia podstawowe cd..

- ⇒ Proces szyfrowania oraz deszyfrowania jest sterowany przez *klucz* lub *klucze kryptograficzne*



Systemy szyfrowania

- Istnieją dwa zasadnicze systemy szyfrowania informacji, tj.:
- ⇒ Szyfry symetryczne lub inaczej z kluczem tajnym
 - ⇒ Systemy niesymetryczne lub inaczej z kluczem jawnym

Cechy systemów z kluczem tajnym

- Bezpieczeństwo algorytmu bazuje na utrzymaniu klucza w ścisłej tajemnicy
- Nadawca i odbiorca muszą uzgodnić klucz przed wymianą informacji
- System nie nadaje się do komunikacji pomiędzy wieloma osobami ze względu na możliwość ujawnienia klucza. (Nie jest tajemnicą informacja, którą zna więcej niż jedna osoba)

Cechy systemów z kluczem jawnym

- ⇒ Wykorzystują różne klucze do szyfrowania i deszyfrowania oraz nie można wyznaczyć w sposób łatwy jednego z nich na podstawie drugiego;
- ⇒ Klucz jawny (szyfrujący) może zostać ujawniony i służy do szyfrowania wiadomości przesyłanych przez dowolne osoby do właściciela klucza jawnego;
- ⇒ Odszyfrowanie wiadomości jest możliwe tylko za pomocą klucza prywatnego;
- ⇒ Klucz prywatny może być wykorzystywany jako podpis elektroniczny. W takim przypadku klucz jawny służy do weryfikowania podpisu.

Rodzaje szyfrów

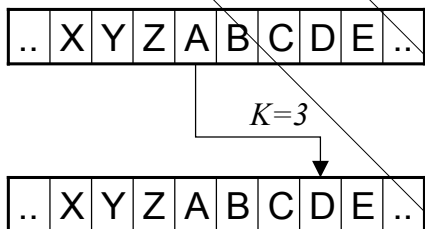
- Szyfry podstawieniowe
 - ⇒ szyfry proste
 - ⇒ szyfry homofoniczne
 - ⇒ wieloalfabetowe
 - ⇒ poligramowe
- Szyfry przestawieniowe
- Szyfry mieszane (podstawieniowo-przestawieniowe)

Szyfry podstawieniowe proste

Przykłady z kryptografii klasycznej

Szyfr Cezara

Książka kodowa



Słowo	Kod
dom	1456
drzewo	5646
wykład	5456
itd...	itd....

Szyfry przestawieniowe

Przykład zapisu wierszowego

Tekst jawny: *matematyka to królowa nauk*

	1	2	3	4	5	6
1	m	a	t	e	m	a
2	t	y	k	a	t	o
3	k	r	ó	l	o	w
4	a	n	a	u	k	x

Zapis tekstu jawnego odbywa się wierszami, a odczyt kolumnami.

Kluczem kryptograficznym jest znajomość kształtu figury i długość wiersza, $K=6$

Tekst zaszyfrowany: *mtkaayrntkóaealumtokaowx*

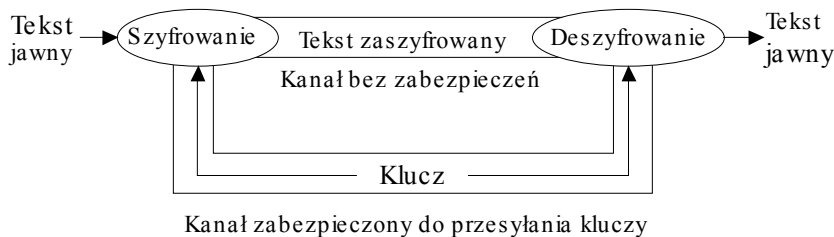
Nowoczesne algorytmy kryptograficzne

- DES (*ang. Data Encryption Standard*)
- AES (*ang. Advanced Encryption Standard*)
- IDEA (*ang. International Data Encryption Algorithm*)
- RSA (*Rivest Shamir i Adleman*)
- DSA (*ang. Digital Signature Algorithm*)
- XOR (*Sumowanie modulo 2 tekstu z kluczem*)

Metody łamania szyfrów

1. Atak bez tekstu jawnego (*ang. ciphertext-only*)
2. Atak ze znanym tekstem jawnym
3. Atak z wybranym tekstem jawnym
4. Atak z adaptacyjnie wybranym tekstem jawnym
5. Atak z wybranym szyfrogramem

Klasyczna ochrona informacji



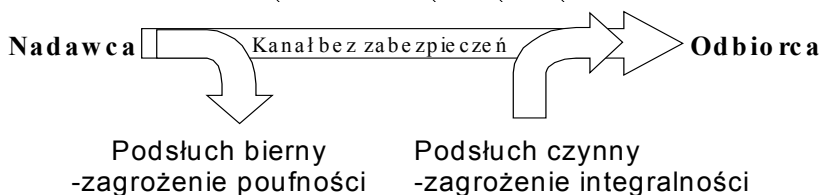
Klucz kryptograficzny przesyłany był kanałem powolnym, ale zabezpieczonym (np. kurierem). Wiadomości i odpowiedzi były przekazywane kanałem narażonym na podsłuch, ale w postaci zaszyfrowanej

Współczesna ochrona informacji

Współczesna kryptografia chroni dane przesyłane kanałami transmisyjnymi lub dane przechowywane w systemach komputerowych.

Współczesna ochrona informacji polega na:

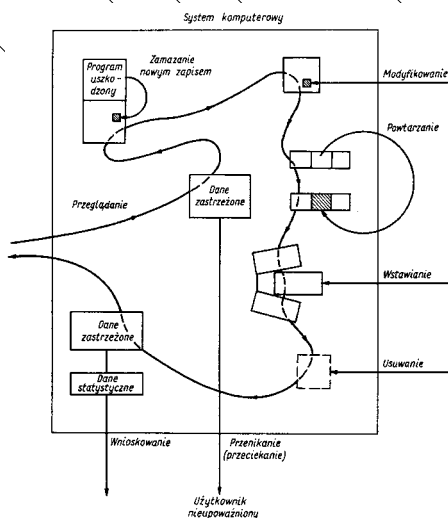
- ⇒ *zapewnieniu poufności (prywatności),*
- ⇒ *zapewnieniu autentyczności (integralności),*
- ⇒ *uwierzytelnianiu*



© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 13/33

Podstawowe zagrożenia dla informacji



- ✓ Przeglądanie
- ✓ Modyfikowanie
- ✓ Zastępowanie
- ✓ Zamazywanie
- ✓ Powtarzanie
- ✓ Wstawianie
- ✓ Usuwanie
- ✓ Wnioskowanie
- ✓ Przenikanie

© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 14/33

Teoria informacji

- Dla dyskretnego źródła informacji miarą *ilości informacji* w wiadomości jest przeciętna liczba bitów niezbędna do zakodowania wszystkich informacji.
- Formalną miarą ilości informacji w wiadomości jest *entropia* tej wiadomości. Mierzy ona nieokreśloność lub nieprzewidywalność informacji. Im większa entropia, tym większa jest ilość informacji zawarta w wiadomości. Zerowa entropia oznacza, że wiadomość nie niesie żadnej informacji.
- *Przykład:* Wiadomość o treści „Ford T, którego kupiliśmy wczoraj jest czarny” nie niesie informacji o kolorze, ponieważ Ford T był produkowany tylko w kolorze czarnym. Przesłana informacja jest z góry zdeterminowana.

Entropia

Entropia jest dana zależnością:

$$H(X) = -\sum_{i=1}^n p(X_i) \cdot \log_2 p(X_i)$$

gdzie: X_1, \dots, X_n będą wariantami treści wiadomości występującymi z prawdopodobieństwami: $p(X_1), \dots, p(X_n)$ oraz:

$$\sum_{i=1}^n p(X_i) = 1$$

Uwzględniając wszystkie wiadomości X , mamy:

$$H(X) = -\sum_X p(X) \log_2 p(X) = \sum_X p(X) \log_2 \left(\frac{1}{p(X)} \right)$$



Entropia warunkowa

Entropia warunkowa *wieloznaczność* jest dana zależnością:

$$H_Y(X) = - \sum_{X,Y} p(X,Y) \cdot \log_2 p_Y(X)$$

Niech Y jest wiadomością ze zbioru Y_1, \dots, Y_n oraz spełnione jest równanie:

$$\sum_{i=1}^n p(Y_i) = 1$$

Prawdopodobieństwo łączne $p(X,Y) = p_Y(X)p(y)$

Prawdopodobieństwo warunkowe $p_Y(X)$

$$H_Y(X) = - \sum_Y p(Y) \sum_X p_Y(X) \log_2 p\left(\frac{1}{p_Y(X)}\right)$$

Optymalne zakodowanie informacji

Entropia osiąga maksymalną wartość, gdy wszystkie informacje są jednakowo prawdopodobne.

$$H(X) = \max, \text{ gdy } p(X_i) = \frac{1}{n}, \quad i = 1, \dots, n$$

Naturalne źródła informacji, nie osiągają maksymalnej entropii ze względu na nierównomierne prawdopodobieństwo występowania zdarzeń. Np. litera a występuje w tekście częściej niż litera x .

Nadmiarowość informacji

- Każde naturalne źródło informacji (tekst, obraz, zapis dźwięku) charakteryzuje się nadmiarowością (redundancją).
- Nadmiarowość zawarta w wiadomości ułatwia złamanie szyfrogramu.
- Dla każdego języka można określić parametry:
 - ⇒ wskaźnik bezwzględny języka R
 - ⇒ wskaźnik względy języka r
 - ⇒ redundancja języka $D=R-r$

Wskaźnik bezwzględny języka R

Wskaźnik bezwzględny języka określa maksymalna liczbę bitów niezbędną do przedstawienia informacji, która mogłaby być zakodowana w dowolnym znaku, przy założeniu, że wszystkie możliwe sekwencje znaków są jednakowo prawdopodobne. Definiowany jest on zależnością:

$$R = \log_2 L,$$

gdzie: L -ilość liter w alfabecie

Wskaźnik względny języka r

Wskaźnik względny języka określa przeciętną liczbę bitów na jeden znak informacji. Definiowany jest zależnością.

$$r = \frac{H(X)}{N \cdot \log_2 L}$$

gdzie: N -ilość znaków w wiadomości.



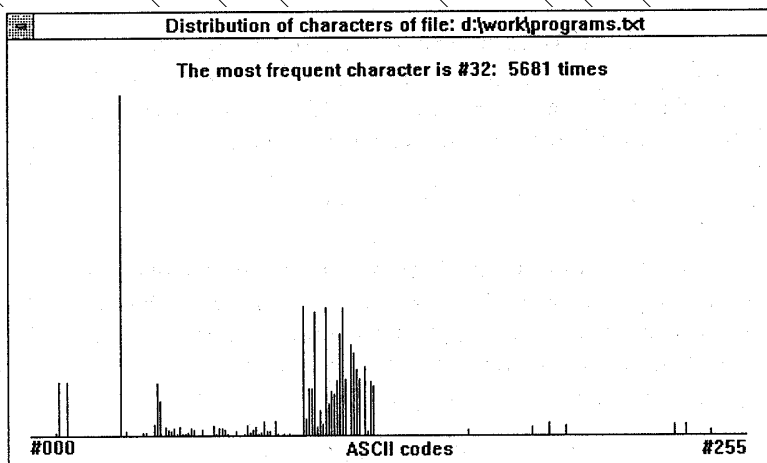
Statystyczne właściwości języka

Statystyczne właściwości języka ułatwiają w znacznym stopniu łamanie szyfrogramów. Dla każdego języka można określić:

- ⇒ Rozkład częstości występowania poszczególnych liter
- ⇒ Rozkład częstości występowania *diagramów* (zlepków dwuliterowych, np.: (ów, rz, ch, itd)
- ⇒ Rozkład częstości występowania *trigramów* (zlepków trzyliterowych, np.: prz, krz, uje



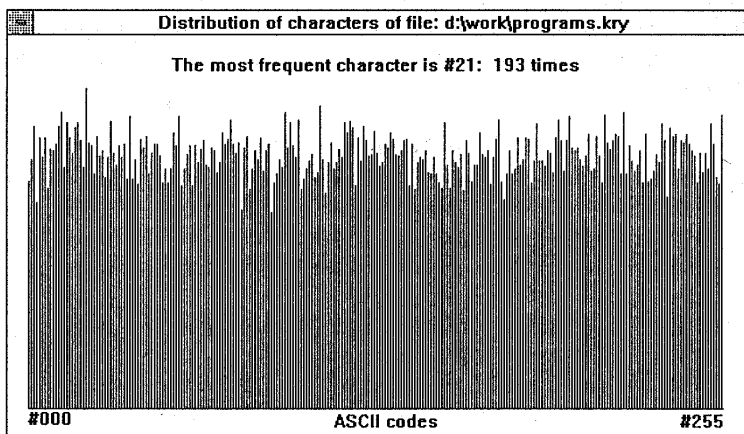
Wykres częstości występowania znaków w tekście



© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 23/33

Wykres częstości występowania znaków po kompresji



© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 24/33

Przejęte oznaczenia

M - wiadomość jawna

C - szyfrogram

K - przestrzeń klucza

E - algorytm szyfrujący

D - algorytm deszyfrujący

A - alfabet źródła

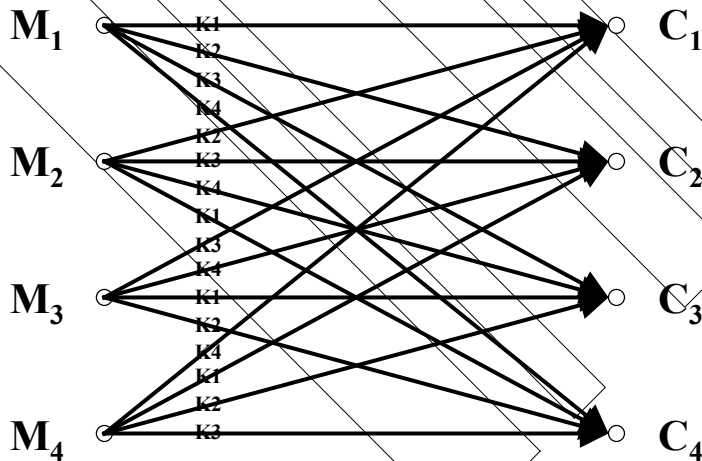
C - alfabet szyfrogramu

m_i - pojedynczy znak
wiadomości jawnej

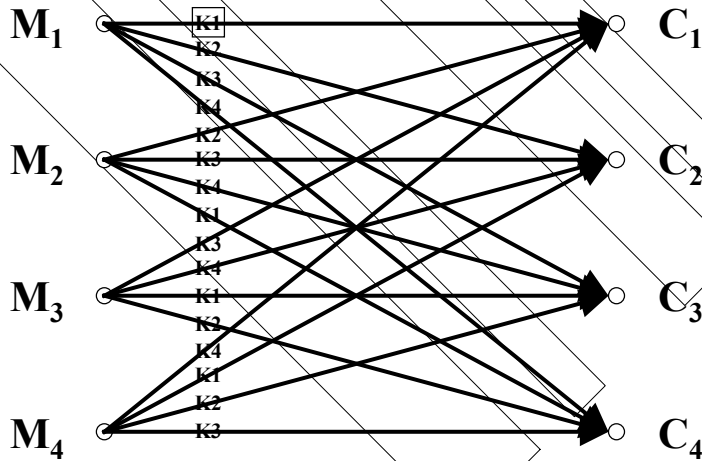
c_i - pojedynczy znak
szyfrogramu

k - klucz

Poufność doskonała



Poufność doskonała



Długość krytyczna kodu

Długość krytyczna jest to najmniejsza długość tekstu zaszyfrowanego liczona w znakach, która jest niezbędna do jednoznacznego określenia klucza. Nie oznacza to jednak, że posiadanie tej ilości informacji umożliwi złamanie kodu.

Długość krytyczną wyznacza się dla warunku

$$H_C(K) = 0$$

gdzie entropia warunkowa klucza jest dana zależnością

$$H_C(K) = \sum_C p(C) \sum_K p_C(K) \cdot \log_2 \left(\frac{1}{p_C(K)} \right)$$

Długość krytyczna kodu cd..

Nie zawsze da się wyznaczyć długość krytyczna kodu. Dla szyfrów *losowych* to jest takich, dla których dla każdego klucza K i szyfrogramu C przekształcenie deszyfrujące jest niezależną zmienną losową o rozkładzie równomiernym na zbiorze wszystkich tekstów jawnych można zapisać, że długość krytyczna N jest równa

$$N = \frac{H(K)}{D}$$

Moc szyfru

Moc szyfru jest określona przez moc obliczeniową potrzebną do jego złamania przy pomocy *algorytmu łamiącego* i mierzona jest przez *złożoność obliczeniową* :

$$O(n) = T(n) \times S(n),$$

gdzie:

T -złożoność czasowa (czas potrzebny do obliczeń)

S -złożoność przestrzenna (niezbędna ilość pamięci)

Zarówno T i S wyrażane są jako funkcje długości n ciągu wejściowego

Klasyfikacja algorytmów

$O(1)$ - Algorytm stały o złożoności niezależnej od długości słowa wejściowego.

$O(n)$ - Algorytm liniowy. Złożoność jest wprost proporcjonalna do długości słowa wejściowego.

$O(n^t)$ - Algorytm wielomianowy. Jego złożoność zależy od n^t przy stałym t .

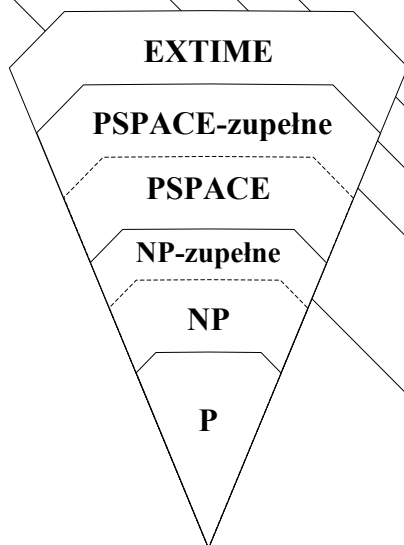
$O(t^{f(n)})$ - Algorytm wykładniczy. t jest stałe, a funkcja $f(n)$ jest wielomianem zmiennej n

Przeważnie złożoność obliczeniowa używana jest do określenia złożoności czasowej algorytmu.

Przykłady złożoności czasowej algorytmów

Klasa	Złożoność	Liczba operacji dla $n = 10^6$	Czas wykonania przy 10^6 operacji/s
Stały	$O(1)$	1	1 μ s
Liniowy	$O(n)$	10^6	1 s
Kwadratowy	$O(n^2)$	10^{12}	11,6 dnia
Sześcienny	$O(n^3)$	10^{18}	32 000 lat
Wykładniczy	$O(2^n)$	10^{301030}	$10^{301006} \times$ wiek istnienia wszechświata

Złożoność problemów



Klasa problemów rozwiązywalnych w czasie wykładniczym za pomocą niedeterministycznej maszyny Turinga

Klasa problemów rozwiązywalnych w przestrzeni wielomianowej ale niekoniecznie w czasie wielomianowym za pomocą niedeterministycznej maszyny Turinga

Klasa problemów rozwiązywalnych w czasie wielomianowym za pomocą niedeterministycznej maszyny Turinga

Klasa problemów rozwiązywalnych w czasie wielomianowym za pomocą deterministycznej maszyny Turinga

© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 33/33

KONIEC

Dziękuję za uwagę

© Robert Borowiec

Kryptografia, Wykład IV
Wprowadzenie, strona 34/33