Kryptografia i bezpieczeństwo danych - wprowadzenie

Sławomir Samolej
ssamolej.kia.prz.edu.pl
ssamolej@prz.edu.pl

1

Literatura

- Kurs internetowy „Cryptography I” https://www.coursera.org,
- Książka w wolnym dostępie: Dan Boneh and Victor Shoup, „A Graduate Course in Applied Cryptography”, 2020, https://toc.cryptobook.us/
- Jonathan Katz, Yehuda Limdell, Introduction to modern cryptography, CRC Press, Third Edition, 2021.

- Jean-Philippe Aumasson, Nowoczesna Kryptografia, Praktyczne wprowadzenie do szyfrowania, PWN 2018
- Marcin Karbowski, Podstawy Kryptografii, Wydanie II, Helion 2014
- Douglas R. Stilson, Maura B. Paterson, Kryptografia w teorii i praktyce (wyd. IV), PWN, 2021.
- Wiliam Stallings, Lawrie Brown, Bezpieczeństwo Systemów Informatycznych, Zasady i Praktyka, Tom 1 i 2, Helion 2019
- David Kahn, Łamacze kodów, historia kryptografii, ZYSK i S-KA, 2019.

2

Plan wykładu

- Wprowadzenie
- Szyfry strumieniowe
- Szyfry blokowe
- Integralność
- Odporność na kolizje
- Szyfrowanie z uwierzytelnieniem
- Wymiana kluczy
- Kryptografia z kluczem publicznym
- Protokoły wymiany kluczy
- Kryptografia z kluczem publicznym
- Aplikacje kryptografii
  - Podpis Elektroniczny
  - VPN
  - …
- Bezpieczeństwo Systemów Informatycznych

3

Kryptografia jest wszędzie…

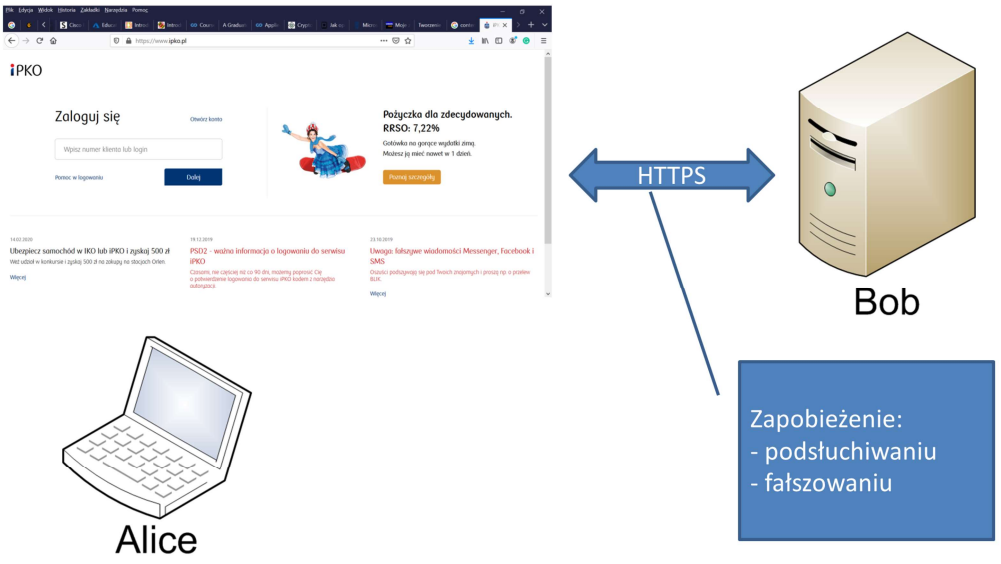
- **Bezpieczna komunikacja**
  - Chroniony dostęp do stron: HTTPS
  - Sieci bezprzewodowe: 802.11i WPA2, GSM, Bluetooth
- **Szyfrowanie danych na dyskach:** EFS, TrueCrypt
- **Ochrona przed tworzeniem nieautoryzowanych kopii:** (np. DVD, Blu-ray): CSS, AACS
- **Uwierzytelnienie użytkowników**

… i wiele innych

4

We współczesnych zastosowaniach informatyki kryptografia jest bardzo często stosowana. Jest stosowana do ochrony komunikacji. Protokół HTTPS chroni nasze połączenie nap. z bankami, czy sklepami internetowymi; nasza łączność bezprzewodowa jest również szyfrowana. W ramach kursu część z wymienionych protokołów będzie szczegółowo omawiana. Kryptografia jest również stosowana do szyfrowania danych na dyskach. Nawet jeśli zostaną one skradzione, to nie bzie można ich odtworzyć. Kryptografia chroni również przed możliwością tworzenia nieautoryzowanych kopii danych (np. DVD, czy Blue-ray) i pozwala weryfikować (autentykować) użytkowników systemu komputerowego.

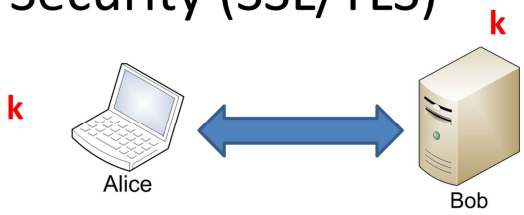
Bezpieczna komunikacja



5

Wracając do chronionej komunikacji, możemy sobie wyobrazić potrzebę wymiany informacji pomiędzy Alice i Bobem. Alice przy pomocy swojego komputera personalnego chce wymienić informację z Bobem – serwerem. Protokół do wymiany informacji nazywa się HTTPS, ale w zasadzie jest to protokół SSL/TLS. Ma on dwa cele zapobieżenie podsłuchania treści komunikatu, ale również zapobieżenie fałszowania treści.

Secure Sockets Layer / Transport Layer Security (SSL/TLS)

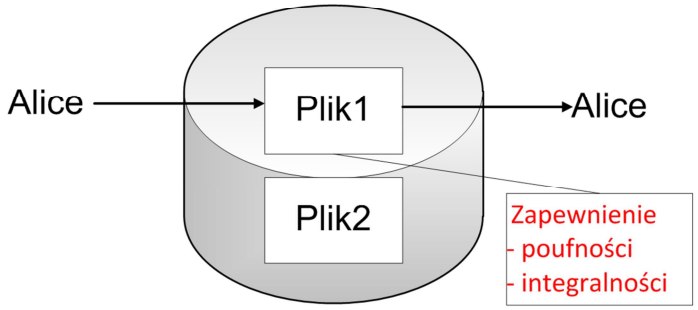


- Dwie podstawowe części:
  - Protokół „handshake” (uzgodnienie): **Ustanowienie współdzielonego tajnego klucza z zastosowaniem kryptografii klucza publicznego**
  - Warstwa „record” (ustalenie formatu przesyłu pakietów): **Transmisja danych z zastosowaniem tajnego klucza symetrycznego**
    Zapewnienie poufności i integralności danych

6

Protokół składa się z dwóch podstawowych etapów: (1) bezpiecznej wymiany kluczy szyfrowania pomiędzy stronami, a następnie (2) wymiany zaszyfrowanych komunikatów w takiej formie, aby nie można było ich odszyfrować oraz podstawić pod nie innych.

Ochrona danych dyskowych

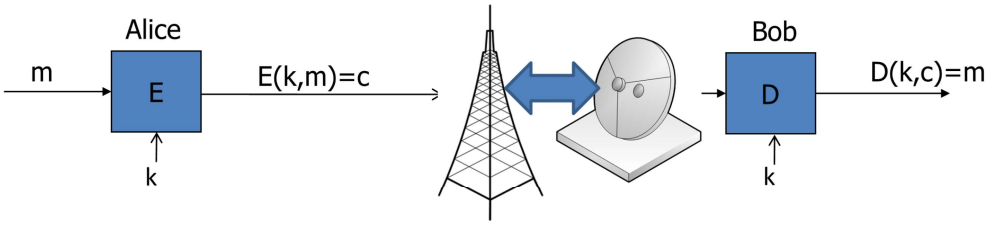


Problem szyfrowania dysków jest analogiczny do chronionej komunikacji, kiedy Alice wysyła wiadomość do Alice

7

Kolejną typową aplikacją kryptografii jest ochrona danych dyskowych. Są tu zdefiniowane dwa cele. Atakujący, który skradnie dysk, nie będzie mógł odczytać zgromadzonych na nim danych. Dodatkowo, jeśli będzie on próbował dane sfałszować, to podczas odszyfrowywania tych danych także zostanie to wykryte. W pewnym sensie szyfrowanie danych na dysku jest podobne do bezpiecznej komunikacji, tylko samego ze sobą. Alice wysyła na dysk zaszyfrowane dane, a po jakimś czasie je odczytuje, czyli przesyła do siebie samej komunikat.

Schemat kryptografii symetrycznej



E, D: algorytm szyfrujący/deszyfrujący
k: tajny symetryczny klucz (np. 128 bitowy)
m, c: wiadomość, zaszyfrowana wiadomość

- Algorytm szyfrowania jest upubliczniony
- Nie używamy własnych szyfrów
- Jedynym tajnym elementem jest klucz

8

Podstawowym blokiem bezpiecznej wymiany informacji jest schemat kryptografii z kluczem symetrycznym. W takim schemacie oboje uczestnicy wymiany informacji współdzielą ten sam tajny klucz, który nie jest znany atakującemu. Będą oni stosować szyfr, który składa się z dwóch algorytmów: E(ang. Encryption) i D (ang. Decryption). Algorytm E (szyfrujący) bierze jako wejście wiadomość m (ang. Message) i klucz k (ang. Key) i wytwarza szyfrogram (zaszyfrowany tekst). Z kolei algorytm D (deszyfrujący) bierze szyfrogram i klucz i odszyfrowuje ukryte dane otrzymując z powrotem wiadomość m.

Bardzo ważnym aspektem współczesnej kryptografii jest fakt, że algorytm szyfrujący jest znany publicznie. Jedynym tajnym elementem systemu jest klucz szyfrowania/deszyfracji. Należy stosować te publicznie znane algorytmy, a nie tajne. Dla tych jawnych algorytmów przeprowadzono wiele analiz i ataków, co podnosi ich wiarygodność.

Tryby użycia klucza

Jednorazowe użycie klucza: (one time key)

- Klucz jest jednorazowo zastosowany do zaszyfrowania wiadomości
  - np. szyfrowanie poczty: nowy klucz jest generowany dla każdej wiadomości

Wielokrotne używanie klucza: (many time key)

- Ten sam klucz jest stosowany do szyfrowania wielu wiadomości
  - np. szyfrowanie plików: ten sam klucz służy do szyfrowania wielu plików
- W takim przypadku będziemy potrzebowali dodatkowych zabiegów, żeby zapewnić bezpieczeństwo

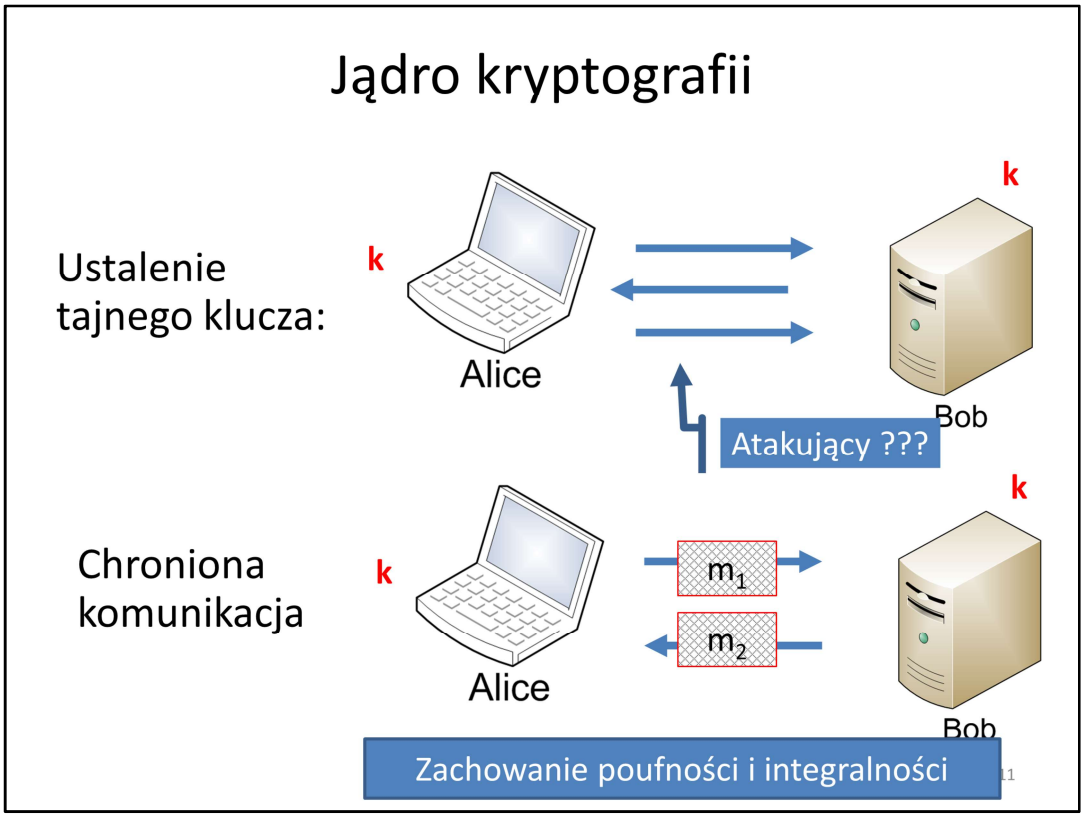
9

W kryptografii symetrycznej mówimy o dwóch sposobach stosowania kluczy. Pierwszy z nich polega na tym, że dla każdej wiadomości stosujemy nowy klucz. Tak może się dziać np. podczas szyfrowania wiadomości poczty elektronicznej. Drugi przypadek (np. gdy chcemy szyfrować dane na dysku) polega na wielokrotnym stosowaniu tego samego klucza. W tym przypadku będziemy musieli stosować dodatkowe zabiegi w celu zabezpieczenia danych.

Uwagi na tym etapie

- Kryptografia jest:
  - Doskonałym narzędziem do ochrony informacji w komputerach
  - Podstawą do stworzenia wielu mechanizmów ochrony przesyłu informacji i przechowywania danych
- Kryptografia nie jest:
  - Rozwiązaniem wszystkich problemów ochrony (np. błędy oprogramowania, ataki stosujące soscjotechniki)
  - Zabezpieczeniem, jeśli jest stosowana w niewłaściwy sposób (np. WEP)
  - Czymś, co trzeba wymyślić samemu
    - Jest bardzo wiele przykładów złamanych systemów szyfrowania powstałych at-hoc
    - Należy się trzymać opublikowanych standardów i schematów kryptograficznych

10

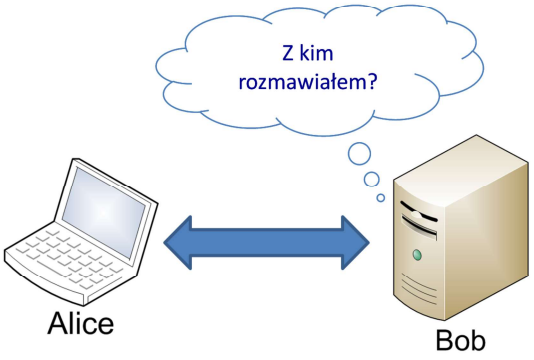


Jeśli chcielibyśmy dokonać przeglądu zastosowań kryptografii, to na pewno pierwszym zastosowaniem jest chroniona komunikacja. Przypomnijmy, że zaczyna się ona od protokołu wymiany tajnego klucza/kluczy. Przy okazji użytkownicy się przedstawiają, czyli Alice wie, że rozmawia z Bobem i Bob wie, że rozmawia z Alice. Potem z zastosowaniem tych kluczy następuje szyfrowanie wiadomości, ale także zapewnienie ich integralności (zabezpieczenie przed fałszowaniem zaszyfrowanych danych).

Kryptografia umożliwia więcej… (1)

- Podpis elektroniczny

- Anonimowa komunikacja (mix net)



12

Podpis elektroniczny jest odpowiednikiem prawdziwego podpisu dokumentu. Podpisując taki dokument gwarantujemy jego treść, to znaczy zakładamy, że nie będzie on mógł być podmieniony na inny (bo na innym nie będzie naszego podpisu), oraz dzięki grafologowi nie będziemy się mogli wyprzeć, że to nie nasz podpis (czyli gwarantujemy, że to my podpisaliśmy ten dokument). W świecie papierowych dokumentów każdy nasz podpis jest taki sam. Niestety nie można tego przenieść do świata cyfrowego. Krótko, stosując kryptografię klucza publicznego i tzw. Infrastrukturę klucza publicznego (certyfikaty) można tak przekształcić wiadomość, że po założeniu pod nią cyfrowego podpisu nie będzie jaj można sfałszować oraz nie będziemy się mogli wyprzeć, że to my podpisaliśmy. Nasz podpis nie będzie mógł być przeniesiony i umieszczony pod innym dokumentem, bo będzie to łatwe do wykrycia.
Innym zastosowaniem kryptografii jest anonimowa komunikacja. Jest to rozwiązanie o nazwie mix-net. Ktoś chce skonsultować się przez Internet w sprawie stanu swojego zdrowia i siłą rzeczy nie jest zainteresowany ujawnieniem swojej tożsamości. Dzięki odpowiednim zastosowaniom metod kryptograficznych możliwe jest zestawienie takiej szyfrowanej komunikacji przez kolejne serwery proxy, zapewniające anonimowość Alice. Wiadomości mogą być przesyłane w obie strony.

Kryptografia umożliwia więcej… (2)

- Anonimowa **cyfrowa** waluta
  - Czy mogę wydać „cyfrową monetę” z zachowaniem anonimowości?
  - Jak zapobiec podwójnemu płaceniu tą samą monetą?

Alice

Internet
anonimowa komunikacja

Kto to był?

13

Anonimowa cyfrowa waluta daje możliwość dokonywania transakcji, tak jak gotówką w realnym świecie. Typowo, podczas zakupu np. książki za gotówkę sprzedawca nie musi znać naszej tożsamości. Celem waluty cyfrowej jest właśnie taka możliwość wymiany towarów w Internecie. Transakcja dochodzi i jest zapisana w bezpiecznych rozproszonych rejestrach (blockchain), ale tożsamość kupującego pozostaje anonimowa. Odpowiednie techniki kryptograficzne zapobiegają również nieuczciwemu mnożeniu pieniędzy przez płacącego. Typowe zakupy przez Internet oznaczają pozostawienie po sobie śladu (karta kredytowa, rachunek bankowy…).

Protokoły

- Wybory

14

Załóżmy, że mamy dwie partie partię 0 i partię 1. Wyborcy głosują na swoją partię. Pewna grupa głosuje na partię 0 inna na partię 1. Generalnie, wygrywa ta partia, która ma większość głosów. Podczas głosowania jesteśmy zainteresowani odczytem stanu głosów, ale niczym więcej (nic więcej nie ma być ujawnione). W elektronicznych wyborach można stworzyć np. cyfrowe centrum wyborcze, które przyjmuje odpowiednio zaszyfrowane głosy, ogłasza wynik wyborów, ale również zabezpiecza przed wielokrotnym wysyłaniem głosów. Wyjściem jest informacja o wynikach wyborów i nic więcej.

Protokoły

- Aukcja w systemie **Vickreya**

15

Podobne zagadnienie dotyczy prywatnych aukcji. Przykładowo, w systemie aukcyjnym Vickera aukcje wygrywa ten, kto zadeklaruje chęć zapłacenia najwyższej stawki, ale na koniec płaci drugą, co do wielkości zadeklarowaną stawkę. Odpowiednio dobrany protokół kryptograficzny ujawni tylko kto wygrał i ile na końcu zostanie zapłacone i nic poza tym.

Protokoły

- Bezpieczne obliczenia wielostronne

Cel: obliczyć $f(x_1, x_2, x_3, x_4)$

Zaufany urząd

"Twierdzenie:" cokolwiek może być wykonane z zastosowaniem zaufanego urzędu może być wykonane bez niego

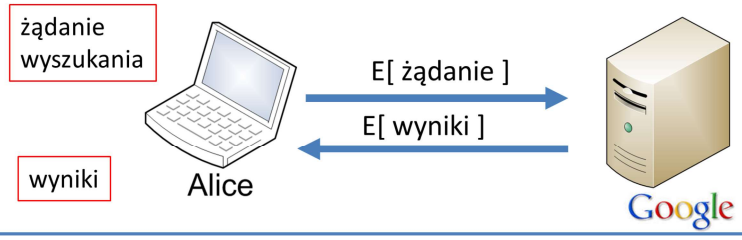
16

Poprzednie przykłady są szczególnymi przypadkami ogólnego zagadnienia zwanego bezpiecznymi obliczeniami wielostronnymi. Każdy z uczestników systemu wprowadza do niego jakiś sekret (ukrytą informację). System ma obliczyć pewną funkcję na danych pobranych z sekretów i ujawnić tylko wynik tej funkcji a nie same przesłane sekrety. Niezbyt bezpiecznym rozwiązaniem jest ustanowienie zaufanego urzędu. Każdy z uczestników obliczeń przesyła swoje dane, obliczenia są przeprowadzone i upublicznione, ale nic więcej nie jest ujawnione, bo mamy umowę z zaufanym urzędem. Problem pojawia się, gdy urząd zostanie skompromitowany…

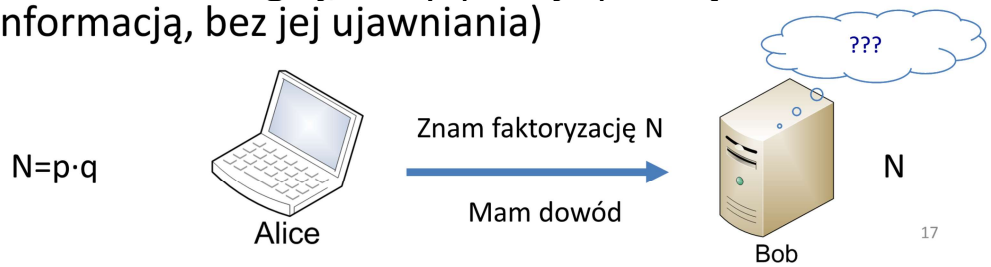
Istnieje jednak twierdzenie, że cokolwiek jesteśmy w stanie policzyć z zastosowaniem takiego zaufanego centrum, jesteśmy w stanie policzyć bez niego.

Magia kryptografii

- Prywatne zlecenia zdalnych obliczeń



- Dowód z wiedzą zerową (jedna ze stron potrafi udowodnić drugiej, że dysponuje pewną informacją, bez jej ujawniania)



17

Precyzyjna nauka

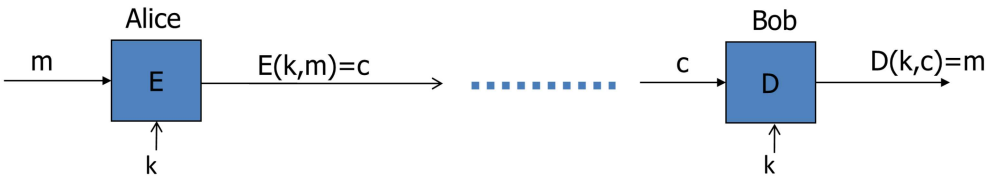
Trzy kroki w kryptografii:

- Precyzyjny model zagrożenia
- Propozycja konstrukcji
- Udowodnienie, że złamanie konstrukcji przy ustalonym zagrożeniu jest bardzo trudnym problemem

Historia kryptografii

19

Szyfry symetryczne



Kilka historycznych przykładów (łatwe do złamania)

- Szyfr podstawieniowy

k :=

```
a -> c
b -> w
c -> n
.
.
.
z -> a
```

21

Szyfr Cezara (bez klucza)
(przesunięcie o 3)

k :=

a -> d
b -> e
c -> f
.
.
y -> b
z -> c

22

Jak złamać szyfr podstawieniowy (1)

Jaka jest liczba możliwych kluczy, przy założeniu 26 liter w alfabecie?

$|K| = 26$

$|K| = 26!$ (silnia)

$|K| = 2^{26}$

$|K| = 26^2$

(liczba wszystkich permutacji)

23

Jak złamać szyfr podstawieniowy (2)

Jaka jest częstotliwość występowania liter w polskim alfabecie?

a	8.91%	w	4.65%	p	3.13%	g	1.42%	ć	0.40%
i	8.21%	s	4.32%	m	2.80%	ę	1.11%	f	0.30%
o	7.75%	t	3.98%	u	2.50%	h	1.08%	ń	0.20%
e	7.66%	c	3.96%	j	2.28%	ą	0.99%	q	0.14%
z	5.64%	y	3.76%	l	2.10%	ó	0.85%	ź	0.06%
n	5.52%	k	3.51%	ł	1.82%	ż	0.83%	v	0.04%
r	4.69%	d	3.25%	b	1.47%	ś	0.66%	x	0.02%

24

Jak złamać szyfr podstawieniowy (3)

1. Należy użyć częstotliwości występowania pojedynczych liter
2. Należy użyć częstotliwości występowania par liter
3. Należy użyć częstotliwości występowania trójek liter

Można zastosować atak z wykorzystaniem tylko zaszyfrowanego tekstu!!!

25

Przykład

(tekst zaszyfrowany w języku angielskim)

UKBYBIPOUZBCUFEEBORUKBYBHOBBRFESPVKBWFOFERVNBCVBZPRUBOFERVNBCVBPCYYFVUFO
FEIKNWFRFIKJNUPWRFIPOUNVNIPUBRNCUKBEFWWFDNCHXCYBOHOPYXPUBNCUBOYNRVNIWN
CPOJIOFHOPZRVFZIXUBORJRUBZRBCHNCBBONCHRJZSFWNVRJRUBZRPCYZPUKBZPUNVPWPCYVF
ZIXUPUNFCPWRVNBCVBRPYYNUNFCPWWJUKBYBIPOUZBCUIPOUNVNIPUBRNCHOPYXPUBNCUB
OYNRVNIWNCPOJIOFHOPZRNCRVNBCUNENVVFZIXUNCHPCYVFZIXUPUNFCPWZPUKBZPUNVR

B	36
N	34
U	33
P	32
C	26

➔ E

➔ T

➔ A

NC	11
PU	10
UB	10
UN	9

digrams

➔ IN

➔ AT

UKB	6
RVN	6
FZI	4

trigrams

➔ THE

26

Szyfr Vigenère'a (16 wiek, Rzym) (przykład z alfabetem łacińskim)

k = C R Y P T O C R Y P T O C R Y P T

(+ mod 26)

m = W H A T A N I C E D A Y T O D A Y

---

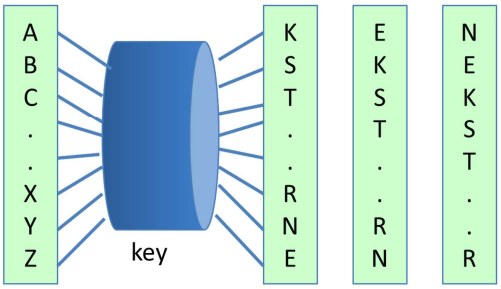
c = Z Z Z J U C L U D T U N W G C Q S

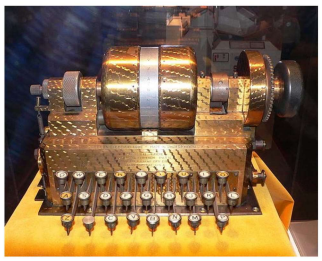
Najczęściej występuje = "H" ⟹ pierwsza litera klucza = "H" – "E" = "C"

27

Maszyny szyfrujące z rotorami (1870-1943)

Wczesny przykład: maszyna Heberna (pojedynczy rotor)

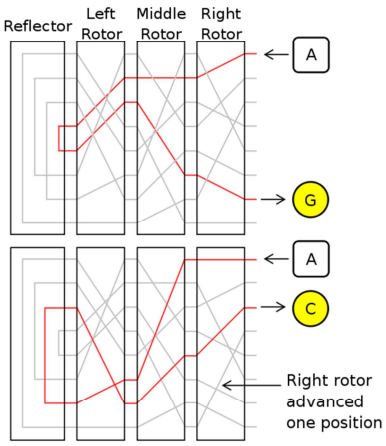




28

Maszyny szyfrujące z rotorami c.d.

Najsłynniejsza: Enigma (3-5 rotorów)



keys = $26^4 = 2^{18}$ (właściwie 2^{36} z powodu łącznicy kablowej)

29

DES – Data Encryption Standard

DES: # keys = 2^{56} , rozmiar bloku = 64 bits

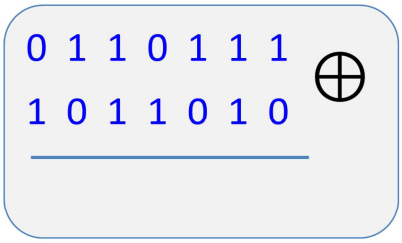
Dziś: AES (2001), Salsa20 (2008) (i wiele innych)

30

Podstawowy aparat matematyczny

Operacja bitowa XOR

Wejście		WYJŚCIE
A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0



31

Ważna właściwość funkcji XOR

- (Twierdzenie, nieformalnie)
  Jeśli klucz szyfrujący jest ciągiem losowym, to wykonanie operacji $\oplus$ (XOR) na dowolnym ciągu bitowym i tym kluczu daje dalej ciągiem losowym.

32

Paradoks dnia urodzin

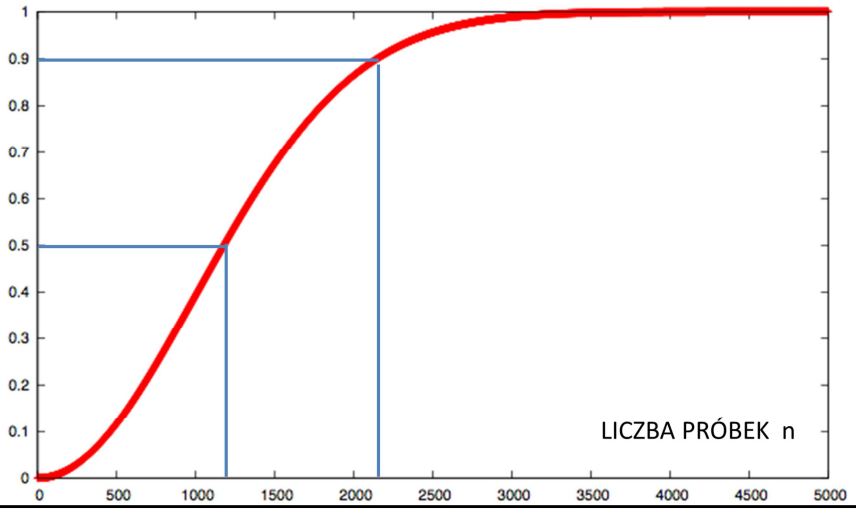
- Problem podstawowy:
  Ile minimalnie osób należy wybrać, żeby prawdopodobieństwo znalezienia wśród nich co najmniej dwóch osób obchodzących urodziny tego samego dnia było większe od 0,5?
  Odpowiedź: 24
- Problem kryptograficzny:
  Jeśli weźmiemy n elementowy zbiór losowych ciągów bitów (mających identyczny rozkład), to analizując $n=1.2\times|U|^{1/2}$ ciągów znajdziemy dwa jednakowe z prawdopodobieństwem ½.

33

Konsekwencja paradoksu dnia urodzin

Generując np. klucze szyfrujące losowe o długości 128 bitów (jest 2^{128} możliwych kombinacji), po sprawdzeniu 2^{64} kluczy dwa z nich z dużym prawdopodobieństwem będą takie same.

$|U|=10^6$



34