

Steganografia

Piotr Turski
Stanisław Witkowski



O czym będzie mowa?

- Wstęp
- Skąd tak głośno o steganografii?
- Krótki ;-) rys historyczny
 - Metody prehistoryczne
 - Metody tekstowe
 - Metody elektroniczne
- Metody komputerowe
- Stegoanaliza
- Znaki wodne
- Podsumowanie

Skąd wziął się pomysł?

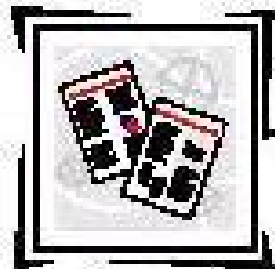
- Czemu po prostu nie szyfrować?
- Co robi przeciwnik, kiedy przechwyci wiadomość?
- Jak rozwiązać ten problem?

Rozwiązanie

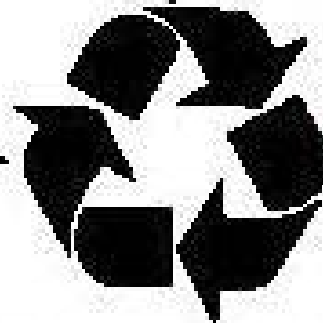
- Chcemy przekazać wiadomość tak, aby sam fakt przesyłania nie został odkryty
- Nośnik wiadomości może być dostępny dla wielu osób, w tym dla przeciwnika – jednak tylko adresat wiadomości będzie wiedział jak odczytać ukrytą treść



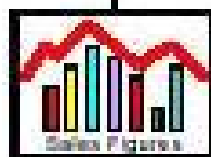
Klucz (stegokey)



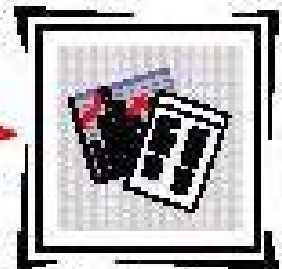
Nośnik - np.
plik graficzny



Program steganograficzny



Wiadomość do ukrycia



Nośnik + ukryta
informacja

Pojęcie „steganografia”

- steganos (gr) – ukryte
- -graphy (pisanie, rysowanie)
- *„Idea steganografii polega na niezauważalnej dla wzroku lub ucha zmianie cyfrowego zapisu tekstu, obrazu lub utworu muzycznego. Owe modyfikacje mogą nieść poufne, praktycznie niemożliwe do odczytania przez osobę niewtajemniczoną, treści.”*

O co tyle hałasu?

- Steganografia istniała w różnych formach od wieków – na światło dzienne została jednak wywleczona w ostatnich latach
- Do momentu masowego korzystania z komputerów i internetu temat ten był dla szarego człowieka całkowicie obcy
- Pierwsza konferencja naukowa dot. Steganografii odbyła się dopiero w 1996r.
- Co się zmieniło?

O co tyle hałasu? c.d.

Z końcem lat 90-tych zaczęto przypuszczać, że steganografią mogą się posługiwać terroryści

Po 11 września nastąpiła w USA masowa histeria – uznawano, że wszędzie dookoła (usenet, ebay, Amazon itd.) są ukryte tajne informacje od Osamy do jego uśpionych agentów, czekających tylko na wezwanie. Terroryści mieli sobie przekazywać informacje o następnych atakach poprzez wymianę plików z ukrytą właściwą treścią (wymiana miałaby następować właśnie poprzez grupy dyskusyjne oraz publiczne chatroomy)

Twierdzono, że ukryta treść jest najczęściej maskowana pod zdjęciami pornograficznymi



A Secret Language

Hijackers May Have Used Secret Internet Messaging Techniques



Bin Laden: Steganography Master?

by [Declan McCullagh](#) | [Also by this reporter](#)

02:00 AM Feb. 07, 2001 PT



WASHINGTON -- If there's one thing the FBI hates more than Osama bin Laden, it's when Osama bin Laden starts using the Internet.

Wireless Hot Spot Directory

Find hot spots near you:

Enter U.S. ZIP code

[Story Tools](#)

[PRINT](#) [MAIL](#)

So it should be no surprise that the feds are getting unusually jittery about what they claim is evidence that bin Laden and his terrorist allies are using message-scrambling techniques to evade law enforcement.

USA Today [reported](#) on Tuesday that

Rys historyczny - początki

- Wszystkie wiadomości były przekazywane przez jednego (zazwyczaj) posłańca
- Dwa sposoby na ukryte przekazanie informacji – Posłaniec niesie w ręku wiadomość fałszywą (przykrywkę), podczas gdy prawdziwą:
 - pamięta
 - ma gdzieś ukrytą
- Przykłady:
 - Chiny
 - Histiaeus z Miletu

Rys historyczny c.d.

- Następny krok – atramenty sympatyczne
 - jakie? Sok z cytryny, mleko, mocz
 - apogeum: I i początki II Wojny Światowej
 - współcześnie: ultrafiolet, kopie dokumentów (VOID)
- Co z kopiowaniem?
 - Przy kopiowaniu przez osobę nieświadomą ukrytej w dokumencie prawdziwej treści, była ona bezpowrotnie tracona

Rys historyczny c.d.

- Johannes Trithemius: Steganographia, ok. 1500r

- Kodowanie:

- Imionami aniołów w spisach

- padiel aporsy mesarpon omeuas peludyn malpreaxo

Rys historyczny c.d.

- Johannes Trithemius: Steganographia, ok. 1500r

- Kodowanie:

- Imionami aniołów w spisach

- padiel aporsy mesarpon omeuas peludyn malpreaxo

- padiel aporsy mesarpon omeuas peludyn malpreaxo

- p r y m u s a p e x

- prymus apex

- Słowami modlitw (szyfr „Ave Maria”)

- Książka tabel litera – słowo

- Każde słowo w modlitwie to słowo z kolejnej tabeli – odczytujemy jaką literę ono oznacza

Łamigłówka ;-)

Dear Business person ; We know you are interested in receiving amazing info ! This is a one time mailing there is no need to request removal if you won't want any more ! This mail is being sent in compliance with Senate bill 2716 , Title 5 , Section 306 . This is not a get rich scheme . Why work for somebody else when you can become rich in 94 WEEKS . Have you ever noticed most everyone has a cellphone and how many people you know are on the Internet ! Well, now is your chance to capitalize on this ! WE will help YOU sell more & increase customer response by 160% . The best thing about our system is that it is absolutely risk free for you ! But don't believe us ! Mr Ames of Illinois tried us and says "My only problem now is where to park all my cars" . This offer is 100% legal . We BESEECH you - act now ! Sign up a friend and you get half off . Thank-you for your serious consideration of our offer . Dear Colleague ; You made the right decision when you signed up for our mailing list . If you are not interested in our publications and wish to be removed from our lists, simply do NOT respond and ignore this mail . This mail is being sent in compliance with Senate bill 1625 ; Title 4 ; Section 304 ! This is not multi-level marketing ! Why work for somebody else when you can become rich in 29 DAYS . Have you ever noticed people love convenience plus most everyone has a cellphone . Well, now is your chance to capitalize on this . We will help you deliver goods right to the customer's doorstep & turn your business into an E-BUSINESS . The best thing about our system is that it is absolutely risk free for you ! But don't believe us . Ms Anderson of Hawaii tried us and says "I was skeptical but it worked for me" ! We are licensed to operate in all states ! We BESEECH you - act now ! Sign up a friend and you get half off ! God Bless

Rozwiązanie łamigłówki

- Umowocześniony pomysł mnicha sprzed dwóch slajdów – zamiast słów modlitw użyto słów ze spamu
- SpamMimic.com
- Cały poprzedni ekran to wynik zakodowania:
„taka sobie wiadomosc”

Rys historyczny c.d.

- II Wojna Światowa:
- Kodowanie wiadomości w samym tekście
- Kodowanie wiadomości w mikrofilmach umieszczanych na tekście

**"Watch out for the dots
- lots and lots of little dots."**

Co po 1945?

- Już w trakcie wojny metody z atramentem sympatycznym okazały się zbyt mało efektywne
- Strach przed powszechnym i praktycznie niewykrywalnym używaniem steganografii w trakcie i po wojnie był tak wielki, że obie strony żelaznej kurtyny wprowadziły liczne ograniczenia prawne
- Masowo kontrolowano nie tylko zawartość międzynarodowej korespondencji, ale również sposób zapisania adresu na kopercie oraz nawet rodzaj i sposób przyklejenia znaczka

Łamigłówka :-)

CO 25 st.

k1, p1 to end; end k1.

k1, p1 to end; end k1.

k1, p1, k across; end p1, k1.

k1, p1, p across; end p1, k1.

k1, p1, k across; end p1, k1.

k1, p10, k3, p10, k1.

k1, p1, k9, p3, k9; end p1, k1.

k1, p10, k3, p10, k1.

k1, p1, k9, p3, k9; end p1, k1.

k1, p4, k15, p4, k1.

k1, p1, k3, p15, k3; end p1, k1.

k1, p4, k15, p4, k1.

k1, p1, k3, p15, k3; end p1, k1.

k1, p5, k13, p5, k1.

k1, p1, k4, p13, k4; end p1, k1.

k1, p5, k13, p5, k1.

k1, p1, k4, p13, k4; end p1, k1.

k1, p6, k11, p6, k1.

k1, p1, k5, p11, k5; end p1, k1.

k1, p6, k11, p6, k1.

k1, p6, k11, p6, k1.

k1, p7, k9, p7, k1.

k1, p1, k6, p9, k6; end p1, k1.

k1, p7, k9, p7, k1.

k1, p1, k6, p9, k6; end p1, k1.

k1, p8, k7, p8, k1.

k1, p1, k7, p7, k7; end, p1, k1.

k1, p8, k7, p8, k1.

k1, p1, k7, p7, k7; end, p1, k1.

k1, p9, k5, p9, k1.

k1, p1, k8, p5, k8; end p1, k1.

k1, p9, k5, p9, k1.

k1, p1, k8, p5, k8; end p1, k1.

k1, p10, k3, p10, k1.

k1, p1, k9, p3, k9; end p1, k1.

k1, p11, k1, p11, k1.

k1, p1, k10, p1, k10; end p1, k1.

k1, p1, p across; end p1, k1.

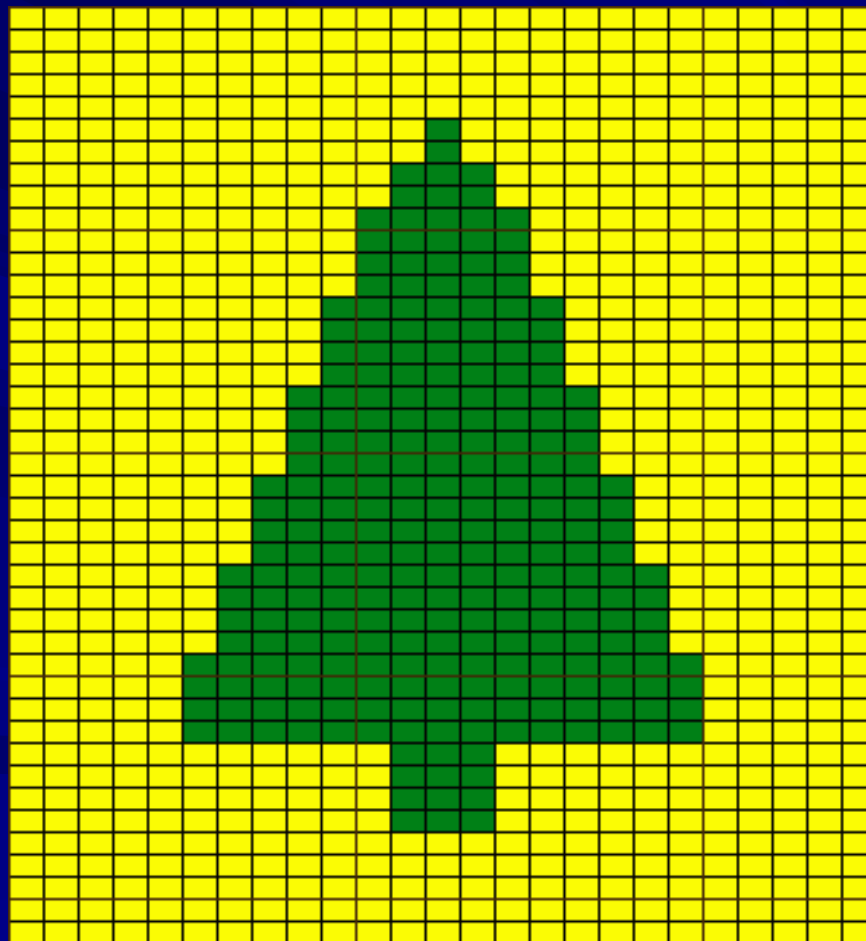
k1, p1, k across; end p1, k1.

k1, p1, p across; end p1, k1.

k1, p1 to end; end k1.

k1, p1 to end; end k1.

Rozwiązanie łamigłówki



Co po 1945? c.d.

- Przykładowe przepisy w USA:
 - Zakaz prowadzenia korespondencyjnych międzynarodowych partii szachowych (przesyłanie kolejnych ruchów na szachownicy)
 - Zakaz przesyłania instrukcji szydełkowania, haftowania, robienia na drutach
 - Zakaz przesyłania wycinków gazet, dziecięcych rysunków
 - całkowity zakaz międzynarodowych zamówień na kwiaty

Co po 1945? c.d.

- Niemiecka metoda z kropkami upowszechniła się
- Z czasem zaczęto stosować mikroprzesunięcia w wydrukach
 - Rozdzielczość drukarki pozwalała na kontrolę wydruku następnej litery nawet do $1/300$ cala
 - Takie kodowanie pozostawało czytelne nawet po wielokrotnym faksowaniu i powielaniu

Metody tekstowe

■ Metoda międzyczdaniowa

- `Cześć! To jest tylko przykład. Kodujemy sobie`
- `raptem parę bitów. I co? I potrzebujemy do`
- `tego aż tylu zdań. Niepraktyczne.`

Metody tekstowe

■ Metoda międzysdaniowa

- Cześć! To jest tylko przykład. Kodujemy sobie
- Cześć! To jest tylko przykład. Kodujemy sobie
- raptem parę bitów. I co? I potrzebujemy do
- raptem parę bitów. I co? I potrzebujemy do
- tego aż tylu zdań. Niepraktyczne.
- tego aż tylu zdań. Niepraktyczne.

Metody tekstowe c.d.

■ Metoda międzyszdaniowa

- Cześć! To jest tylko przykład. Kodujemy sobie
- Cześć! 1To jest tylko przykład.0Kodujemy sobie
- raptem parę bitów. I co? I potrzebujemy do
- raptem parę bitów.0I co? 1I potrzebujemy do
- tego aż tylu zdań. Niepraktyczne.
- tego aż tylu zdań. 1Niepraktyczne.

10011

Metody tekstowe

■ Metoda końca linii

■ Przykładowa pierwsza linijka.

■ I druga linijka – też długa.

■ Ostatnia, krótsza.

Metody tekstowe

■ Metoda końca linii

- Przykładowa pierwsza linijka.
- Przykładowa pierwsza linijka.__ 2
- I druga linijka - też długa.
- I druga linijka - też długa.===== 5
- Ostatnia, krótsza.
- Ostatnia, krótsza._ 1

0 spacji = 000, 1 = 001, 2 = 010, 3 = 011,
4 = 100, 5 = 101, 6 = 110, 7 = 111

Metody tekstowe c.d.

- After the theater, all clients keep
- a tab down at Wesley's Nook.

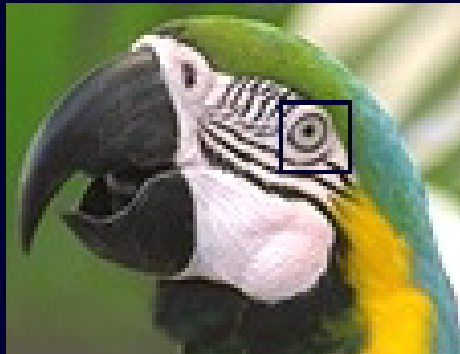
Metody tekstowe c.d.

- After the theater, all clients keep
- After the theater, all clients keep
- A t t a c k
- a tab down at Wesley's Nook.
- a tab down at Wesley's Nook.
- a t d a w N
- ATTACK AT DAWN

Metody tekstowe c.d.

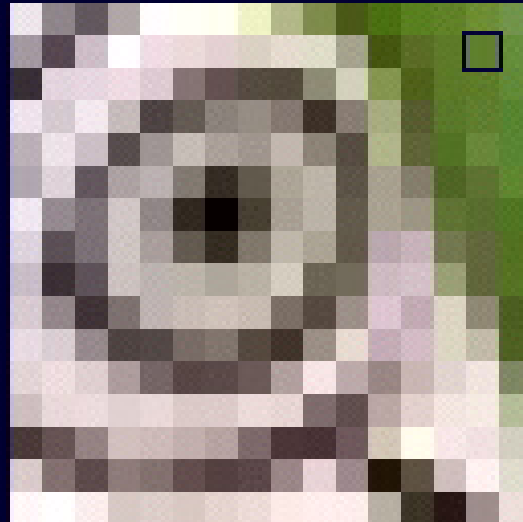
- Metoda syntaktyczna:
- Niektóre zbiory słów brzmią dobrze w każdej kolejności:
 - Listy (imion, zakupów, marek towarów itd.)
- Każda możliwa kombinacja słów oznacza jedną kombinację bitów:
- „Drogi mężu, kup drożdże, cukier, mąkę i ogórki” 001
- ... drożdże, cukier, ogórki i mąkę” 010
- ... drożdże, mąkę, cukier i ogórki” 011
-

Metody komputerowe



Parrot

Pixels



97: 01100001
128: 10000000
45: 00101101

Red-Green-Blue values

Grafika

LSB

■ Least Significant Bit

10010101	00001101	11001001
10010110	00001111	11001010
10011111	00010000	11001011

Kodujemy 101101101

10010101	0000110 0	11001001
1001011 1	0000111 0	1100101 1
10011111	00010000	11001011

Formaty

■ Skala szarości

Kodowanie 256 odcieni. Ludzkie oko rozróżnia ok. 64

■ True Color

Wrażliwość ludzkiego oka: R:G:B = 3:6:1
(najbardziej wrażliwe na zmiany zieleni)

pixel RGB = 00011010; 01010101; 00011110

kodujemy: 01000001

wynik: 00011001; 01010100; 00000001

Paleta kolorów

■ Paleta kolorów

		3	

obraz

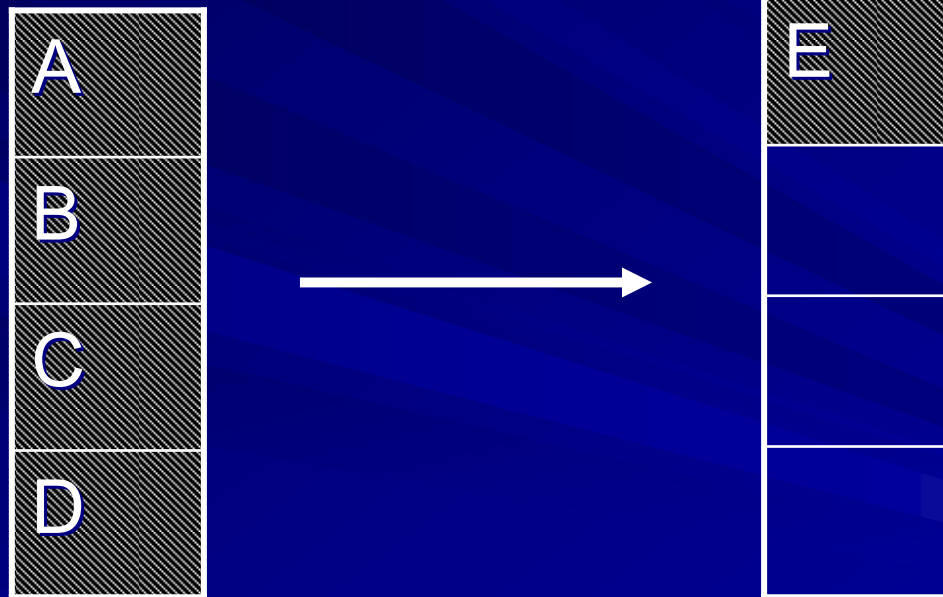


	R	G	B
1			
2			
3	105	243	18
4			
5			

paleta

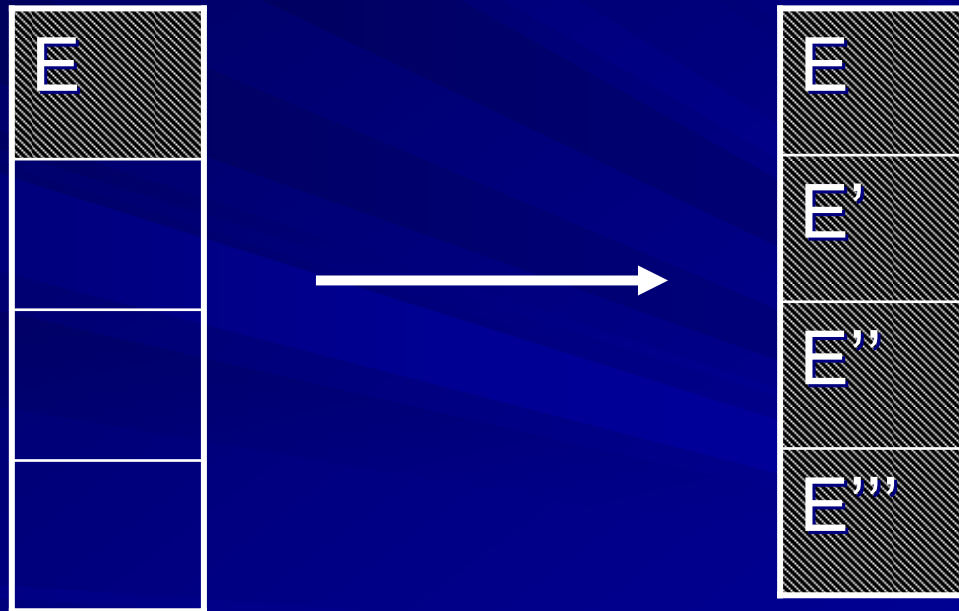
Paleta kolorów

■ 1. Redukcja palety



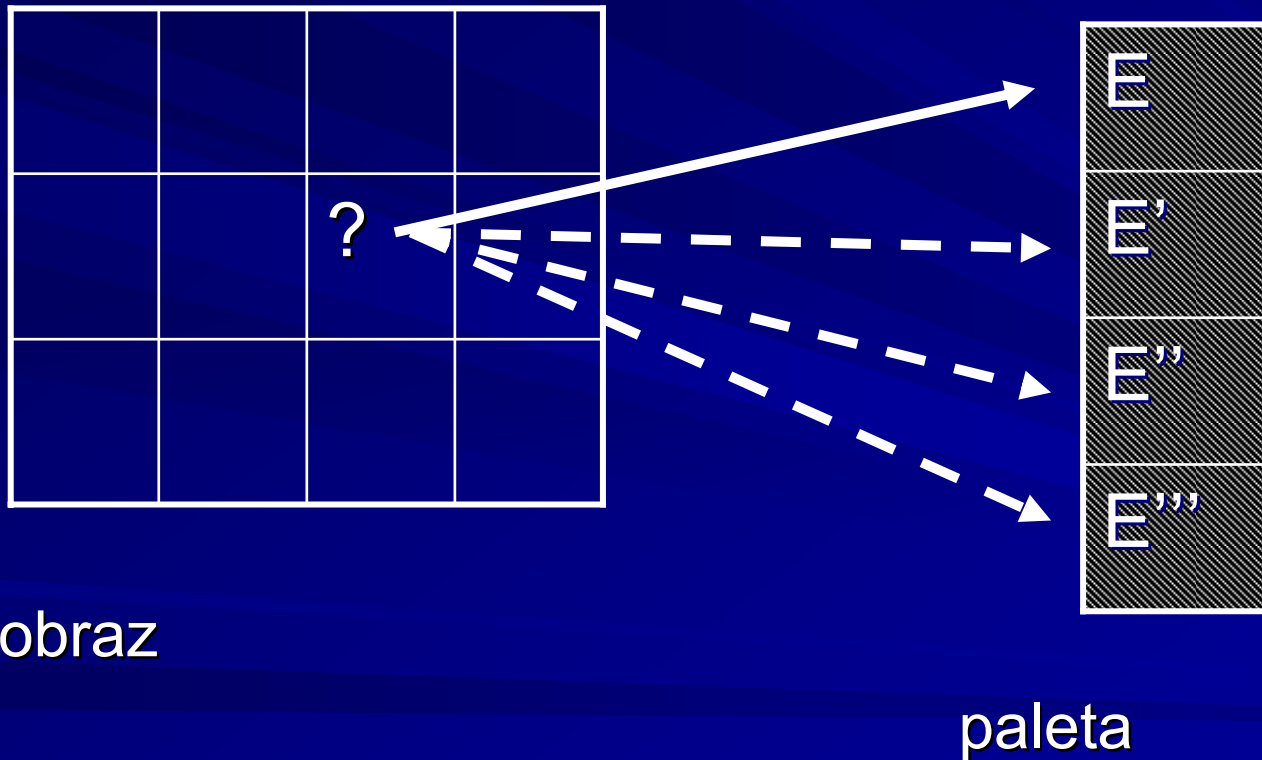
Paleta kolorów

■ 2. Powielenie palety



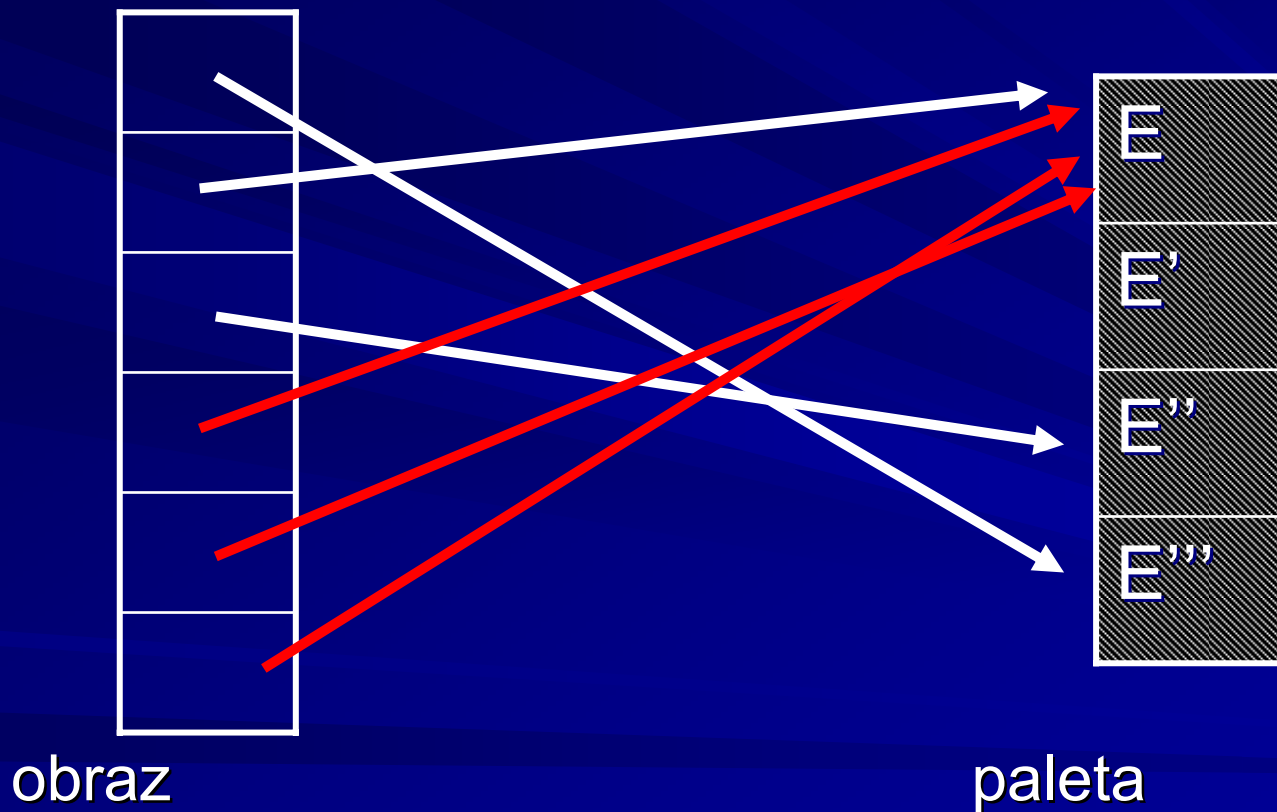
Paleta kolorów

■ 3. Właściwe kodowanie



Paleta kolorów

- 4. Dodanie szumu w indeksach poza tajnymi danymi



Paleta kolorów

■ 5. Wymieszanie palety

A	B	C	D	E
A	B	C	D	E
A	B	C	D	E
A	B	C	D	E



A	B	C	D	E
C	A	B	E	D
B	D	E	C	A
A	D	C	E	B

Kompresja stratna

- Ukrywanie informacji w istotnych elementach obrazu
- Odporność na przypadkowe lub celowe uszkodzenie obrazu

metody typu Transform Domain Techniques:

- dyskretna transformacja cosinusowa (DCT)
- transformacja falowa

- Metody niezależne od formatu pliku

Wybór nośnika

- Dużo szczegółów (zdjęcia)
- Bez płynnych przejść między kolorami
- Dużo większy niż kodowana informacja

Steganografia w TCP/IP

- Pierwsze pomysły: pingowanie maszyny
- Wiadomości w polach nagłówka nieużywanych w praktyce lub tych które powinny mieć konkretną wartość
problem: firewall'e zerują te pola
- Początkowe numery sekwencyjne (ISN)
używane przy nawiązywaniu połączenia
z założenia generowane losowo
problem: co 4 bajty potrzebne nowe połączenie
wygląda jak skanowanie portów

TCP/IP cd...

- Pomysł: timestamp
pole używane do optymalizowania prędkości transmisji
nikt po drodze nie zmienia tej wartości
w wielu systemach rozdzielczość rzędu 10ms
potraktujmy ostatni bit jako losowy
- Znając klucz można transmitować niewykrywalne wiadomości publicznie znanym algorytmem
- Dokładne analizy wykazały że bit nie jest losowy
- Poprawka: podsłuchać rozkład i dostroić generator

Sieci

- Html – tagi, wielkość liter, nazwy załączników...
- CRC – celowe niszczenie ramki
takie ramki są standardowo odrzucane przez aplikacje

Stegoanaliza

- *„Stegoanaliza jest sztuką zapobiegania lub wykrywania steganografii. Polega ona z reguły na usunięciu ukrytej informacji lub sprowadzeniu jej do bezużyteczności.”*
- The Three D's of Defeating Steganography:
 - Detect
 - Decrypt
 - Destroy

Detekcja

- sprawdzanie plików graficznych pod kątem nieregularności
 - Szukanie bardzo bliskich sobie par pikseli
 - Sprawdzanie liczby kolorów
- Sprawdzanie daty ostatnich modyfikacji
- Różnice w używanych formatach na stronach www

Detekcja

- Stegdetect – University of Michigan
 - Analizuje częstotliwość DCT (Discrete Cosine Transform) ;-)
- Securestego – Professor Jessica Fridrich, State University of New York, Binghamton
 - Analiza par kolorów w obrazku – szukanie znacznego wzrostu praktycznie identycznych par

Skasowanie

- Przechwycić wiadomość
- Zmienić ją
 - Przyciąć do mniejszego formatu
 - Zmienić głębie kolorów
 - Zachować w innym formacie (gif,jpg,bmp)
 - Zwiększyć stopień kompresji (stratnej)
 - Dowolna kombinacja powyższych
- Zapisać i przesłać dalej jako oryginalnie otrzymaną
- Ukryty przekaz zostanie prawdopodobnie (choć nie na pewno!) zniszczony

Deszyfrowanie

- W założeniu bardzo proste – cytat:
- *Beg, borrow, buy or steal:*
 - *The password*
 - *The location of the message*
 - *The encryption type*
 - *The software used*
- *And then open the message at your leisure*

Deszyfrowanie

- Potrzebny oryginał – prostsze, mało praktyczne
- Niepotrzebny oryginał – trzeba znać wspólny klucz, posiadać dobry generator liczb pseudolosowych

Odporność na ataki

- Problem w ocenie odporności steganografii (lepiej wierzyć praktykom)
- Ukrycie kolejnych wiadomości
- Freeware'owy StegDetect (XSteg) – wskazuje JPEG'i z ukrytą informacją.
 - rzadko wykrywane wiadomości poniżej 1% objętości nośnika
 - powyżej 10%: wszystkie

Czy ktoś tego wszystkiego używa?

- Terroryści ?
- Twórcy (znaki wodne)
- Programy przeszukujące sieć najczęściej potencjalnie ukrytych informacji wykryły na aukcjach, stronach pornograficznych, czatach sportowych
- Łatwe dla każdego (darmowy ScramDisk3 ukrywa całe systemy plików)

Słabości steganografii

- Wiadomości muszą pozostać względnie małe
 - Większość obrazków w internecie jest niewielkich rozmiarów (<15KB)
 - Plik wielkości 1-2KB (strona tekstu) może być z łatwością ukryty przy wykorzystaniu steganografii)
 - Większy rozmiar ukrywanych danych może znacznie powiększyć początkowy rozmiar pliku – przykrywkę, co zwiększa prawdopodobieństwo wykrycia (zbyt duży rozmiar w stosunku do innych podobnych plików na stronie budzi podejrzenia)
- W plikach audio liczba możliwych do upchnięcia danych jest nawet mniejsza (maksymalnie 600B w pliku 10KB)

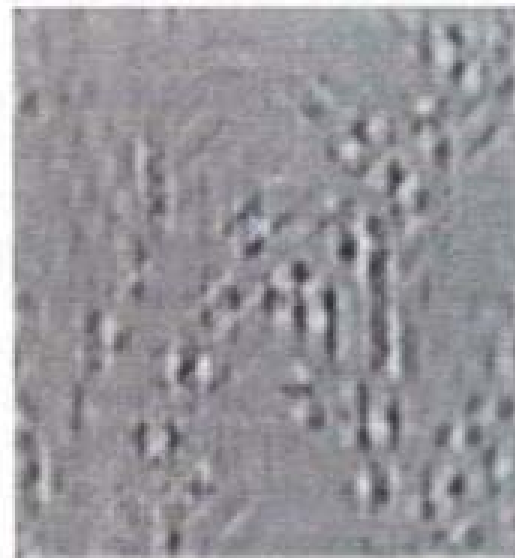
Znaki wodne



Oryginal



Znakowany wodnie



Znak wodny

Znaki wodne

- Badanie głównie nad niewidzialnymi znakami wodnymi
- Ulotne
- Odporne (kompresja, rozmywanie, zmiana rozmiaru, kontrastu, jasności)

Jak to jest z tymi terrorystami?

- W sierpniu 2001 University of Michigan przeprowadził kompleksowe badania witryny ebay:
 - Sprawdzono ponad dwa miliony plików graficznych
 - 17,000 uznano za podejrzane pod względem steganografii
 - 15,000 mogło zawierać ukryty inny obraz
 - Próby złamania haseł dostępowych do tych plików (o ile rzeczywiście pliki te zawierały jakąś ukrytą treść zabezpieczoną hasłem) spełzły na niczym.
 - Testowy plik został poprawnie znaleziony, wykryty, a hasło zostało złamane
 - Podobne przeszukiwanie miało miejsce w usenecie, z podobnym skutkiem
- Więcej informacji: www.citi.umich.edu/u/provos

Słabości badania UoM

- Szukanie jedynie pewnego zakresu treści, tylko w plikach jpg
- Sprawdzanie plików przygotowanych programami Outguess, Jsteg, JPHide
- Brak testów pod kątem programu S-Tools (dość popularnego)
- Brak testów plików .bmp, .gif, .wav
- Terroryści mogli teoretycznie używać całkowicie własnych aplikacji steganograficznych
- Poszukiwania „jedynie” w dwóch miejscach: eBay, usenet
- Słownikowe metody łamania haseł korzystały jedynie ze słowników: angielskiego, niemieckiego, francuskiego oraz zbioru popularnych słów z Koranu oraz literatury popularnej
- Ewentualne hasła mogły być zwyczajnie odporne na metody słownikowe (hasła długie, zawierające cyfry i znaki niealfanumeryczne)

Wnioski UoM

■ Trzy możliwości:

- Użytkownicy internetu nie korzystają w zasadzie ze steganografii
- Nikt nie używa programów, których działania poszukiwano
- Wszyscy użytkownicy programów steganograficznych umieją wybrać odpowiednio skomplikowane hasła)

Ciekawostka



Wybrane linki

■ Narzędzia:

www.jjtc.com/Steganography/toolmatrix.htm – 146 narzędzi

■ „projekty hakerskie”:

„Diggin Em Walls (part 3) – Advanced/Other Techniques for ByPassing Firewalls”

<http://neworder.box.sk/user.php?name=Ka0ticSH>

Phrack magazine 49, 1996-11-08

www.phrack.org/show.php?p=49

„Firewall bypass via protpcol steganography”

www.networkpenetration.com/protocol_steg.html