Akademia Techniczno-Humanistyczna
w Bielsku-Białej

Wydział Budowy Maszyn i Informatyki

Laboratorium z sieci komputerowych

Ćwiczenie numer:
10

Temat ćwiczenia:

Systemy szyfrowania informacji.

1. Wstęp teoretyczny.

Od początku istnienia sieci komputerowych przechowywanie informacji jest narażone na różne niebezpieczeństwa. Wraz z rozwojem techniki komputerowej rosną zagrożenia przechowywanych zbiorów danych i informacji.

Ogólnie mówiąc niebezpieczeństwa czyhające na nasze zasoby możemy podzielić na cztery grupy:
Utrata zasobów - nieodwracalne zniszczenie części lub całości informacji. Utrata cennych danych, będących niejednokrotnie wynikiem długotrwałej pracy i przedstawiających poważną wartość rynkowa, może być rezultatem działań ludzkich lub załamania się systemu.

Kradzież zasobów - przywłaszczenie danych, stanowiących własność innych osób. Kradzież zasobów wiąże się z poważnym zagrożeniem wtedy, gdy złodziej jest w stanie poprawnie je zinterpretować i wykorzystać. Na domiar złego kradzież zasobów nie zawsze wiąże się z utrata danych (dlatego poszkodowany często o jej fakcie nie ma pojęcia).

Przekłamanie zasobów - świadoma lub nieświadoma zmiana części informacji w bazie danych, zazwyczaj związana z chęcią osiągnięcia zysku dzięki fałszerstwu. Dane zwykle bywają zmieniane w minimalnym zakresie, i jeżeli dokonano tego w dyskretny sposób, jest to niesłychanie trudne do wykrycia.

Łamanie praw autorskich dotyczących programów - modyfikowanie i nielegalne kopiowanie programów komputerowych. W przypadku firm zajmujących się wytwarzaniem oprogramowania wiąże się to często z ogromnymi stratami finansowymi.

Szyfrowanie to, jak łatwo się domyśleć, sposób ochrony informacji przed jej zinterpretowaniem przez osoby niepowołane. Mogą one ją odczytywać, lecz zaszyfrowana treść (kryptogram) nie stanowi dla nich żadnej wartości, gdyż nie da się go przekształcić na tekst jawny (otwarty) bez znajomości odpowiedniego klucza.

Najczęściej utajnianymi informacjami są dane personalne. Często zdarza się również, ze autor lub użytkownik systemu potrzebuje zamieścić w nim, bądź tez w zbiorach pomocniczych pewne cechy charakterystyczne, mające służyć do porównywania właściwego środowiska programu ze stanem aktualnym. Takimi informacjami są najczęściej: hasła, procedury uwierzytelniania, numer seryjny programu, nazwa programu, informacje o producencie, długość programu, data jego powstania, sumy kontrolne, niektóre informacje o środowisku komputera (np. fragmenty BIOS-u), cechy kluczy (dyskietek kluczowych czy tez kluczy sprzętowych).

Dziedzinę wiedzy i badań zajmującą się utajnionym zapisywaniem danych nazywamy kryptografia, zaś termin kryptoanaliza obejmuje dziedzinę wiedzy badającą metody przełamywania szyfrów.

Rozróżniamy dwa podstawowe rodzaje szyfrów: przestawieniowe i podstawieniowe.

Szyfry przestawieniowe

Szyfry te zmieniają uporządkowanie bitów lub znaków w danych według pewnego schematu. Zazwyczaj dokonuje się przestawienia za pomocą pewnej figury geometrycznej. Szyfrowanie przebiega wiec w dwóch krokach: tekst jawny wpisuje się do figury w sposób określony pewną tzw. ścieżką zapisu, a następnie odczytuje się go według określonego porządku (ścieżki odczytu) otrzymując tekst zaszyfrowany. Klucz obejmuje więc figurę geometryczną oraz ścieżki zapisu i odczytu.

Pierwszym przykładem szyfru przestawieniowego jest prosty szyfr płotowy. Litery tekstu jawnego zapisuje się tu tak, aby tworzyły kształt przypominający wierzchołek płotu zbudowanego ze sztachet. Tekst zaszyfrowany otrzymujemy odczytując kolejne wiersze tak utworzonej konstrukcji. Poniżej przykład:

Tekst jawny	**T**	**E**	**K**	**S**	**T**	**N**	**I**	**E**	**Z**	**A**	**S**	**Z**	**Y**	**F**	**R**	**O**	**W**	**A**	**N**	**Y**
Klucz k = 3	**T**				**T**				**Z**				**Y**				**W**			
		E		**S**		**N**		**E**		**A**		**Z**		**F**		**O**		**A**		**Y**
			K				**I**				**S**				**R**				**N**	
Tekst zaszyfrowany	**K**	**I**	**S**	**R**	**N**	**E**	**S**	**N**	**E**	**A**	**Z**	**F**	**O**	**A**	**Y**	**T**	**T**	**Z**	**Y**	**W**

Bardzo często używaną figurą geometryczną jest macierz dwuwymiarowa. Jako przykład szyfru weźmy tzw. przestawienie kolumnowe. Tekst jawny zapisuje się do macierzy wierszami. Kryptogram powstaje jako odczyt kolumn w określonym porządku. Poniżej przykład:

Tekst jawny: **TEKSTPRZEDZASZYFROWANIEM**

Ustalamy kolejność komun (np. **1-4-2-5-3**)

T	**E**	**K**	**S**	**T**
P	**R**	**Z**	**E**	**D**
Z	**A**	**S**	**Z**	**Y**
F	**R**	**O**	**W**	**A**
N	**I**	**E**	**M**	**_**

Zatem:

Tekst zaszyfrowany: **TPZFNSEZWMERARITDYA_KZSOE**

Kryptoanalitycy mogą łatwo rozpoznać, czy zastosowany szyfr jest szyfrem przestawieniowym, ponieważ częstość występowania liter tekstu zaszyfrowanego będzie zbliżona do częstości ich występowania w tekście jawnym. Dlatego właśnie tego rodzaju szyfry mogą być w prosty sposób łamane metodą anagramową, polegającą na odtworzeniu właściwej kolejności przemieszanego zestawu znaków.

Szyfry podstawieniowe

W szyfrach podstawieniowych zastępuje się bity, znaku lub bloki znaków odpowiednimi zamiennikami. Istnieją cztery typy szyfrów podstawieniowych:

a) monoalfabetyczne
b) homofoniczne
c) wieloalfabetyczne
d) poligramowe

Szyfry monograficzne

W szyfrach monograficznych każdy znak tekstu jawnego zostaje zamieniony na odpowiedni znak kryptogramu, przy czym w całej wiadomości do zamiany każdego znaku jawnego na zaszyfrowany stosuje się odwzorowanie typu jeden do jednego.

Najbardziej znanym przykładem szyfru monograficznego jest prosty szyfr Cezara (jako pierwszy użył go Juliusz Cezar). Polega on na przyporządkowaniu każdej literze alfabetu odpowiedniego numeru identyfikacyjnego (np. A=0, B=1 itd.) i dokonaniu przesunięcia numeru każdej litery tekstu jawnego o k - pozycji (ma tu miejsce tzw. przewijanie - gdy kończy się alfabet przesuwamy się do jego początku). Zakres szyfrowania można oczywiście rozszerzyć na zbiór znaków ASCI lub jakiś inny skończony zbiór n znaków. Funkcja szyfrująca będzie się wówczas wyrażała wzorem:

$$E(a) = (a + k) \bmod n$$

Przykład szyfu Cezara poniżej:

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Tekst jawny: **TEKSTPRZEDZASZYFROWANIEM**

Przyjmujemy klucz k = 3.

Zatem T + 3 = W itd.

Tekst zaszyfrowany: **WHNVWSUCHGCDVCBIURZDQLHP**

W niektórych szyfrach podstawieniowych monoalfabetycznych do kodowania służyły również niestandardowe alfabety szyfrowe. Przykładem może tu być szyfr, w którym zastosowano zamianę liter na nuty.

Niestety szyfry monoalfabetyczne są w prosty sposób łamane na podstawie analizy częstości występowania liter lub znaków.

Szyfr homofoniczny

Szyfry homofoniczne, podobnie jak poprzednio opisane szyfry monoalfabetyczne, zamieniają każdy znak tekstu jawnego na odpowiedni znak kryptogramu, z tą jednak różnicą, że odwzorowanie ma tu charakter jeden do wielu i każdy znak może być zaszyfrowany jako jeden z pewnej grupy znaków alfabetu szyfrowego.

Przykładem szyfru homofonicznego może być prosty szyfr, w którym litery alfabetu są szyfrowane jako liczby całkowite z przedziału (0, 99), przy czym ilość liczb całkowitych przydzielonych danej literze jest proporcjonalna do względnej częstości jej występowania i żadna z tych liczb nie jest przydzielona do więcej niż jednej litery.
Poniżej przykład:

Tekst jawny: **TEKSTPRZEDZASZYFROWANIEM**

Homofony:

A	**01,35,28,59,82,**	**N**	**15,57,**
D	**48,58,**	**O**	**40,47,66,77,**
E	**27,69,72,87,**	**P**	**24,79,**
F	**37,60,**	**R**	**42,68,94**
G	**06,71,**	**S**	**12,55,97**
I	**08, 31,88,99**	**T**	**22,50,67,92**
J	**29,70,**	**W**	**52,78,**
K	**32,54,64,74,**	**Y**	**39,80,**
M	**04,62,**	**Z**	**19,51,65,75,85,**

Tekst zaszyfrowany: **502774129224687569481982558580609466522815087262**

Szyfry homofoniczne mogą być znacznie trudniejsze do złamania, gdy liczba homofonów przydzielona danej literze jest proporcjonalna do częstości jej występowania w tekście, który chcemy zaszyfrować, ponieważ rozkład częstości występowania symboli jest wtedy prawie jednostajny, co utrudnia analizę. Dodatkową zaletą tych szyfrów jest możliwość kodowania równolegle z autentyczną wiadomością, którą chce się przekazać, wiadomości fałszywej (np. poprzez stosowanie wartości nieprzypisanych żadnemu znakowi).

Szyfry wieloalfabetowe

W szyfrach wieloalfabetowych stosuje się wiele odwzorowań znaków tekstu jawnego na znaki kryptogramu, przy czym każde odwzorowanie jest z reguły typu jeden do jednego, podobnie jak w szyfrach monoalfabetycznych. Jak wiec łatwo zauważyć szyfry wieloalfabetyczne ukrywają rozkład częstości przez użycie wielu podstawień.

Większość szyfrów tej grupy to szyfry okresowe o okresie d znaków. Klasycznym przykładem może tu być powstały w XVI wieku szyfr Vigenere'a.

Szyfrowanie wiadomości przebiega tu na podstawie dowolnie wybranego słowa kluczowego (hasła). W przypadku znaków ASCI może to być dowolny ich ciąg. Do numeru każdego kolejnego znaku tekstu jawnego dodajemy numer odpowiadającego mu znaku słowa kluczowego i uzyskujemy znak kryptogramu. Gdy słowo kluczowe się skończy, bierzemy je kolejny raz od początku. Dla znaków ASCI szyfr Vigenere'a można przedstawić za pomocą poniższej funkcji:

$$F_i(a) = (a + k_i) \bmod 255$$

Poniżej przykład szyfru Vigenere'a dla alfabetu łacińskiego:

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Tekst jawny: **TEKSTJAWNY**

Ustalamy słowo szyfrujące: **RADIO**

Tekst jawny	T	E	K	S	T	J	A	W	N	Y
Wartość	19	4	10	18	19	9	0	22	13	24
Hasło	R	A	D	I	O	R	A	D	I	O
Wartość	17	0	3	8	14	17	0	3	8	14
Wartość	10	4	13	0	7	0	0	25	21	12
Tekst zaszyfrowany	K	E	N	A	H	A	A	Z	V	M

Tekst zaszyfrowany**: KENAHAAZVM**

Jak łatwo zauważyć, że im dłuższe i bardziej skomplikowane jest hasło, tym trudniej odszyfrować tekst utajniony. Z kolei równie łatwo jest zauważyć, że gdy nasze hasło będzie jednoznakowe otrzymamy prosty szyfr monoalfabetyczny.

Szyfry poligramowe.

Szyfry przestawieniowe i podstawieniowe szyfrują krokowo po jednej literze tekstu jawnego. Szyfry poligramowe szyfrują w jednym kroku większe grupy liter i to właśnie powoduje, że złamanie takiego szyfru jest dużo trudniejsze, a to dzięki zachwianiu równowagi pomiędzy częstotliwością występowania liter w tekście jawnym i zaszyfrowanym.

Jednym z szyfrów poligramowych jest szyfr Playfaira, który jest diagramowym szyfrem podstawieniowym. Szyfr ten był stosowany przez Anglików w czasie pierwszej wojny światowej. Kluczem jest macierz o wymiarach 5x5 składająca się z liter (bez litery J).

H	**A**	**R**	**P**	**S**
I	**C**	**O**	**D**	**B**
E	**F**	**G**	**K**	**L**
M	**N**	**Q**	**T**	**U**
V	**W**	**X**	**Y**	**Z**

Przebieg procesu szyfrowania:

Każdą parę liter tekstu jawnego m_1m_2 szyfruje się według podanych reguł (c_1, c_2 - to znaki szyfrogramu)

1. Jeśli litery m_1 i m_2 są w tym samym wierszu, to c_1 i c_2 są znakami położonymi z prawej strony m_1 i m_2;
2. Jeśli litery m_1 i m_2 znajdują się w tej samej kolumnie, to c_1 i c_2 są znakami położonymi poniżej m_1 i m_2;
3. Jeżeli m_1 i m_2 znajdują się w różnych wierszach i kolumnach, to c_1 i c_2 są brane z przeciwległych rogów prostokąta wyznaczonego przez m_1 i m_2, przy czym c_1 pochodzi z wiersza zawierającego m_1, a c_2 z wiersza zawierającego m_2.
4. Jeśli $m_1 = m_2$, to do tekstu jawnego między te litery wstawia się nieznaczącą literę (np. V), co eliminuje powtórzenia.
5. Jeśli tekst jawny ma nieparzystą liczbę znaków, to na końcu tekstu jawnego dopisuje się nieznaczącą literę.

Pierwszą kolumnę macierzy traktuje się jako położoną na prawo od ostatniej kolumny, a pierwszy wiersz jako leżący pod ostatnim wierszem.

Poniżej przestawiono przykład użyciu szyfru Playfaira:

Tekst jawny: **TEKSTPRZEDZASZYFROWANIEM**

Dla pierwszych dwóch znaków: m_1 = **T**, m_2 = **E**

H	**A**	**R**	**P**	**S**
I	**C**	**O**	**D**	**B**
E	**F**	**G**	**K**	**L**
M	**N**	**Q**	**T**	**U**
V	**W**	**X**	**Y**	**Z**

Zatem c_1 = **M**, c_2 = **K**

Tekst zaszyfrowany: **MK LP YD SX KI WS BS WK OG AC MC MV**

2. Plan wykonania ćwiczenia

1. Napisać program umożliwiający szyfrowanie plików tekstowych przy użyciu dowolnego szyfru przestawieniowego.

2. Napisać program umożliwiający szyfrowanie plików tekstowych przy użyciu szyfru homofonicznego. Utworzyć własną tabelę homofonów.

3. Napisać program, korzystający z szyfru Playfaira szyfrujący pliki tekstowe.

3. Literatura.

1. **http://schranz.art.pl/marcin/irc-trawa1/security/wstep.htm**

2. **Dorothy Elizabeth, Robling Denning - Kryptografia i ochrona danych.**

3. **Andrzej Kierzkowski - Ochrona programow i danych w praktyce.**
   Gliwice, Helion, 1992

4. **http://stud.wsi.edu.pl/~sismolna/szyfry.html**