INŻYNIERIA BEZPIECZEŃSTWA – LABORATORIUM NR 1

WSTĘP DO KRYPTOGRAFII

Wstęp

Wyraz *kryptologia* ma swoje starogreckie korzenie, oznaczające „*ukryty”* i „*słowo”*. Kryptologia jest ogólną nazwą nauki, zajmującej się problemami tajnego porozumiewania się. Kryptologię można podzielić na dwie części: kryptografię i kryptoanalizę.

1. Kryptografia i kryptoanaliza

Pod pojęciem *kryptografii* należy rozumieć każdą metodę zapisu informacji, za pomocą tajnego kodu lub szyfru. Obecnie społeczeństwa są w wielkim stopniu zależne od komputerów, przetwarzających ogromne ilości informacji, które przesyła się w sieciach komputerowych i przechowuje w bazach danych. Większość tych informacji powinna być chroniona przed ich odczytaniem przez nieupoważnionych użytkowników systemów i sieci komputerowych, zwłaszcza użytkowników nielegalnych, a kryptografia jest jedynym ekonomicznym środkiem, zapewniającym taką ochronę.

Tak wiec kryptograf poszukuje metod dla zapewniania wiadomościom tajności i potwierdzania ich autentyczności, podczas gdy zadaniem kryptoanalityka jest łamanie opracowanych przez kryptografa szyfrów oraz fałszowanie szyfrowanych wiadomości w taki sposób, by wyglądały na autentyczne.

2. Kontrowersje związane z powszechnym stosowaniem kryptografii

Kryptografię do niedawna stosowano prawie wyłącznie w dziedzinie wojskowości i w dyplomacji. Dlatego tez kryptolodzy pracowali zwykle pod nadzorem organów, związanych z bezpieczeństwem państwa, a wszystkie prace naukowe na temat kryptografii i kryptoanalizy utajniano. Dopiero pod koniec lat 60-tych wielka grupa uczonych różnych narodowości, nie kontrolowana przez rodzime służby specjalne, zainteresowała się problemami kryptologii, publikując znaczną liczbę prac naukowych z tej dziedziny, dzięki czemu kryptograficzna ochrona danych znalazła szerokie zastosowania cywilne.

Stosowanie metod kryptograficznych jest jednak w wielu krajach zabronione. Istnieją państwa, nawet uważane za bardzo demokratyczne, które podejmują próby kontrolowania systemów kryptograficznych, nie używanych przez wojsko czy dyplomację. Np. w USĄ forsowano stosowanie specjalnego układu szyfrującego o nazwie *Clipper*, który generował kryptogramy z łatwością łamane przez amerykańskie służby specjalne. Francja i Rosja zakazały używania metod kryptograficznych dla celów prywatnych. Szukając uzasadnienia dla takich zachowań należy uwzględnić fakt, że kryptografia gwarantuje ochronę informacji przed ich używaniem przez osoby i instytucje do tego

nieupoważnione. Niektórzy właściciele danych mogą traktować policje i organa ścigania właśnie jako instytucje, przed którymi informacje należy ukrywać. Ma to miejsce wówczas, gdy metody elektronicznego przetwarzania danych używane są przez organizacje przestępcze lub działające na peryferiach prawa. Z drugiej strony, powszechnie dostępne sieci komputerowe używane są nie tylko przez uczciwych obywateli, co może stwarzać poważne zagrożenia dla państw demokratycznych, które w najprostszym przypadku starają się zabronić używania środków kryptograficznych lub ograniczyć ich stosowanie. Chodzi tu o narzucanie stosowania takich metod kryptograficznych, które posiadają tzw. tajne przejścia (ang. *backdoor*), umożliwiające służbom bezpieczeństwa deszyfrowanie kryptogramów bez znajomości klucza.

Argumenty te nie przekonują jednak wielu rozsądnych ludzi, bowiem organizacje przestępcze bez skrupułów używają technik i metod kryptograficznych dla swoich celów wbrew wszelkim zakazom. Uczciwa część społeczeństw, respektująca ustalone w tym zakresie zakazy, nie może więc chronić swoich danych przed kryminalistami, jest zatem ewidentnie dyskryminowana. Nie jest tajemnicą, że plagą dla użytkowników sieci są włamania komputerowe, przed którymi chronią najskuteczniej właśnie metody kryptograficzne. Dlatego też środowiska gospodarcze lansują pogląd, że stosowanie kryptografii jest niezbędne, ponieważ podstawowym warunkiem światowego rozwoju gospodarczego jest rozbudowa globalnej sieci komputerowej, stanowiącej infrastrukturę informacyjną dla gospodarki. W tym kontekście uważa się, że ograniczenia w stosowaniu kryptografii mogą spowodować spadek atrakcyjności danego kraju dla inwestorów zagranicznych. Na rynku pojawiło się wiec wiele produktów kryptograficznych, a część z nich dopuszczono w niektórych krajach do powszechnego użytkowania.

3. Zasady konstruowania mocnych systemów kryptograficznych

W trakcie procedury szyfrowania wiadomość oryginalna, czyli **tekst jawny**, zamienia się w wiadomość zaszyfrowaną, zwaną też najczęściej **kryptogramem**. Wejściem dla algorytmu szyfrującego jest wiec tekst jawny oraz tajny klucz, a wyjściem jest kryptogram. Procedura deszyfrowania stosuje algorytm, dla którego danymi wejściowymi są klucz i kryptogram, a wynikiem działania algorytmu deszyfrującego jest tekst jawny. Podstawą każdego projektu systemu kryptograficznego, składającego się z algorytmów szyfrowania i deszyfrowania, jest ocena skuteczności ochrony danych przed nieupoważnionym dostępem, jaka projektowany system potrafi zapewnić. Ustalono niezbyt precyzyjnie, ze system kryptograficzny uważa się za mocny, gdy nakład pracy, wymagany do złamania wygenerowanych za pomocą tego systemu kryptogramów, jest dostatecznie wielki. W trakcie wieloletnich badan opracowano następujące warunki, jakie powinien spełniać mocny algorytm kryptograficzny:

- złożoność matematyczna algorytmu szyfrującego powinna wykluczać możliwość stosowania metod analitycznych do złamania szyfru;
- koszt lub czas, wymagany dla uzyskania klucza lub odczytania tekstu jawnego z kryptogramu powinien być znaczny i praktycznie nie do przyjęcia.

Algorytm szyfrujący powinien spełniać powyższe warunki nawet wówczas, gdy wrogi kryptoanalityk posiada dostęp do względnie dużej porcji tekstów jawnych i odpowiadających im kryptogramów oraz gdy zna on dokładnie wszelkie detale

algorytmu. Zakłada się tez, ze kryptoanalityk dysponuje dużymi i szybkimi komputerami. Biorąc pod uwagę powyższe założenia, można stwierdzić, że bezpieczeństwo mocnego systemu kryptograficznego polega na tajności klucza. Mocne algorytmy kryptograficzne, odpowiadające powyższej definicji, mogą teoretycznie zostać złamane, jednak w praktyce zdarza się to niezmiernie rzadko.

4. Wskazówki Shannona dotyczace konstruowania mocnych systemów kryptograficznych

Claude Elmwood Shannon, twórca teorii informacji, jest jednym z wybitnych matematyków amerykańskich, którzy byli zatrudnieni w czasach II wojny światowej przez wojskowe agendy rządowe USA jako doradcy z dziedziny kryptologii. W aktualnej do dziś pracy „The Communication Theory of Secrecy Systems", opublikowanej w czasopiśmie Bell System Technical Journal w 1949 roku, sformułował on zasady konstruowania systemów kryptograficznych, które generują trudne do złamania szyfry. Definiując model matematyczny bezpiecznego, inaczej mówiąc mocnego systemu kryptograficznego, Shannon wyszedł z założenia, że kryptoanalityk powinien na podstawie kryptogramu wyznaczyć klucz i/lub tekst jawny. Pomóc może mu w tym informacja o pewnych parametrach statystycznych tekstu jawnego (np. częstotliwość występowania znaków), na podstawie której można określić, czy tekstem jawnym jest program, napisany w języku C, fragment prozy w języku japońskim, czy też plik dźwiękowy. W każdym z takich tekstów jawnych jest pewien nadmiar informacyjny, ułatwiający kryptoanalizę. Opracowano zatem wiele testów statystycznych, opartych na teorii informacji, które efektywnie pomagają kryptoanalitykom łamać szyfry, jeśli parametry statystyczne tekstu jawnego są znane. Takie testy, opracowane jeszcze latach 40-tych ubiegłego wieku, są do dziś utajnione.

Według Shannona system kryptograficzny, zapewniający zabezpieczenie doskonałe przed nieupoważnionym dostępem nie może dostarczać żadnej informacji statystycznej o tekście jawnym. Dowodząc tej tezy Shannon wykazał, że ma to miejsce wówczas, gdy liczba kluczy kryptograficznych jest co najmniej tak duża, jak liczba możliwych tekstów jawnych. Klucz powinien więc mieć tyle samo lub więcej bitów, znaków lub bajtów co tekst jawny i żadnego klucza nie powinno się użyć więcej niż jednokrotnie. Ten zaproponowany przez Shannona system kryptograficzny nazywa się systemem szyfrowania z kluczem jednorazowym (ang. o*ne-time-pad*).

Aby można było zredukować nadmiar informacyjny tekstu jawnego i nie przenosić tego nadmiaru do kryptogramu, Shannon zaproponował technikę mieszania i rozpraszania, co w praktyce sprowadza się do naprzemiennego stosowania blokowych szyfrów przestawieniowych i podstawieniowych.

5. Przykład blokowego szyfru przestawieniowego

Zasada konstrukcji takiego szyfru zostanie wyjaśniona na prostym przykładzie, gdzie tekst jawny i kryptogram zapisuje się za pomocą nastepującego alfabetu o 32 znakach:

ABCDEFGHIJKLMNOPQRSTUVWXYZ . : , ; -

składającego się z 26 dużych liter alfabetu łacińskiego, spacji, kropki, dwukropka, średnika i myślnika.

Zakłada się, że blok tekstu jawnego będzie zawierał 48 znaków powyższego alfabetu, rozmieszczonych w 6 wierszach.

0	1	1	2	3	3	0	3
1	3	2	2	3	0	0	3
3	0	3	2	0	1	0	3
2	1	0	2	1	2	3	2
1	2	1	1	0	3	1	2
2	1	0	1	0	0	2	3

Tabela 1. Tablica do konstrukcji szablonów szyfrujacych

Przy tych założeniach można zastosować następujący algorytm szyfrujący dla szyfru przestawieniowego:

1. Utworzyć tablicę o 6 wierszach i 8 kolumnach, wypełnioną losowo liczbami 0, 1, 2 i 3, w której liczba zer, jedynek, dwójek i trójek jest taka sama i wynosi 12. Należy się przy tym starać, aby było jak najmniej sąsiadujących ze sobą komórek, zawierających takie same liczby. Powyższe warunki spełnia tablica, pokazana w tabeli 1, którą otrzymano za pomocą następującego programu, napisanego w Pascalu:

```
var i, k, r: Byte;
l: array[0..3] of Byte;
begin
      repeat
            Randomize;
            for i:=0 to 3 do l[i]:=0;
                  for i:=1 to 6 do begin
                        for k:=1 to 8 do begin
                              r:=Random(4); Inc(l[r]);
                              Write(r:2);
                        end;
                  WriteLn
                  end;
            WriteLn;
      until (l[0]=l[1]) and (l[0]=l[2]) and(l[0]=l[3])
end.
```

2. Wyciąć cztery prostokąty z papieru, o kształcie tablicy z kroku 1. i utworzyć z nich cztery szablony o numerach 0, 1, 2 i 3. W każdym szablonie należy wyciąć 12 otworów w miejscach komórek, zawierających liczbę, oznaczającą numer szablonu.

3. Przygotować czysty i pełny kawałek papieru o wymiarze szablonu, na którym będzie zapisany kryptogram. Następnie do tego kawałka papieru przykładać kolejno przygotowane w kroku 2. szablony, wpisując do wyciętych w szablonach okienek 48 liter tekstu jawnego (4 razy po 12 liter).

A	D	I	U	N	K	T	-
L	E	C	T	U	R	E	R
A	F	O	R	Y	Z	M	-
A	P	H	O	R	I	S	M
A	K	T	O	R	K	A	-
A	C	T	R	E	S	S	.

Tabela 2. Blok tekstu jawnego

A	U	R	A	R	K	D	A
E	-	P	H	A	I	U	C
T	N	R	O	K	R	T	E
R	A	-	I	F	S	S	M
O	A	R	Y	L	S	Z	K
T	M	E	-	C	T	O	.

Tabela 3. Blok kryptogramu szyfru przestawieniowego

Algorytm deszyfrujacy jest prostszy i składa się tylko z jednego kroku:

1. Przykładać kolejno do kryptogramu szablony i w wyciętych okienkach odczytywać kolejne znaki tekstu jawnego, zapisując je jednocześnie w formacie kryptogramu, tj. w sześciu wierszach i ośmiu kolumnach.

Przypuśćmy, ze tekstem jawnym jest fragment słownika polsko-angielskiego, pokazany w tabeli 2. Po zaszyfrowaniu tego tekstu zgodnie z podanym algorytmem otrzyma się kryptogram jak w tabeli 3. Można łatwo sprawdzić, ze z tego kryptogramu, po zdeszyfrowaniu według przedstawionego wyżej algorytmu, otrzymuje się z łatwością właściwy tekst jawny.

1	13	14	25	37	38	2	39
15	40	26	27	41	3	4	42
43	5	44	28	6	16	7	45
29	17	8	30	18	31	46	32
19	33	20	21	9	47	22	34
35	23	10	24	11	12	36	48

Tabela 4. Permutacja szyfrująca

1	7	14	15	18	21	23	27
37	43	45	46	2	3	9	22
26	29	33	35	36	39	42	44
4	11	12	20	25	28	30	32
34	40	41	47	5	6	8	10
13	16	17	19	24	31	38	48

Tabela 5. Permutacja deszyfrująca

Opisany w taki sposób blokowy szyfr przestawieniowy, którego kluczem są szablony szyfrujące i kolejność ich użycia w algorytmie szyfrowania, można też przedstawić w sposób bardziej zmatematyzowany. Jeśli 48 znaków bloku tekstu jawnego ponumeruje się zgodnie z kolejnością ich zapisu, to wynikający z zastosowania szablonów proces szyfrowania przyjmie postać tablicy, pokazanej w tabeli 4, z której wynika, że pierwszym znakiem kryptogramu jest pierwszy znak tekstu jawnego, drugim znak

trzynasty, trzecim czternasty itd. Za pomocą tej tablicy, która reprezentuje pewną permutację zbioru liczb {1, 2, . . . , 48}, można zatem bardziej dokładnie opisać procedurę szyfrowania, niż za pomocą szablonów szyfrujących. Co więcej, traktując te permutacje jako klucz szyfrujący, można precyzyjnie określić liczbę wszystkich możliwych kluczy szyfrujących dla powyższego szyfru, która wynosi oczywiście

$$48! = 12413915592536072670862289047373375038521486354677760000000000,$$

czyli w przybliżeniu $1{,}241391559 \times 10^{62}$. Nie jest to liczba mała, skoro liczbę atomów na naszej planecie szacuje się na 10^{51}.

Do deszyfrowania kryptogramów omawianego kodu należy więc zastosować permutację odwrotną do permutacji szyfrującej, pokazaną w tabeli 5. Na pierwszy rzut oka wydawałoby się, że omawiany szyfr można złamać, próbując po kolei 48! permutacji. Zakładając, ze potrafimy w ciągu jednej sekundy przetestować milion permutacji (co jest założeniem raczej mocno optymistycznym, nawet biorąc pod uwagę obecny stan technologii), czas złamania jednego kryptogramu zająłby około 10^{47} lat. Tymczasem całkowity czas istnienia wszechświata ocenia się na 10^{11} lat. Jeśli jednak do złamania szyfru kryptoanalitycy zastosują znane tylko im testy statystyczne, to z takim szyfrem potrafią się często uporać w ciągu kilku minut.

6. Przykład blokowego szyfru podstawieniowego

W najprostszym przypadku konstruowanie blokowych szyfrów podstawieniowych polega na systematycznym zastępowaniu każdego znaku tekstu jawnego innym znakiem. Algorytm szyfrowania musi wiec stosować tablicę, w której zapisana jest reguła tego podstawiania, a taka tablica jest zapisem pewnej permutacji alfabetu, używanego do reprezentacji tekstu jawnego i kryptogramu. Do deszyfrowania stosuje się oczywiście permutacje odwrotna.

Przyjmijmy, że w rozważanym przypadku używa się tego samego alfabetu, co w przykładzie szyfru przestawieniowego, że 48-znakowy blok tekstu jawnego jest taki sam, jak poprzednio, i że podczas szyfrowania stosuje się tablice podstawień jak w tabeli 6:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
X	H	Y	U	Z	R	.	V	Q	A	:	W	T	,	B	C
Q	R	S	T	U	V	W	X	Y	Z		.	:	,	;	-
S	D	;	E	I	G	L	F	J	K	-	M	P	N	0	

Tabela 6. Tablica znaków tekstu jawnego i odpowiadających im znaków kryptogramu

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
J	O	P	R	T	X	V	B	U	Y	Z	W	.	,	;	:
Q	R	S	T	U	V	W	X	Y	Z		.	:	,	;	-
I	F	Q	M	D	H	L	A	C	E	-	G	K	N	S	

Tabela 7. Tablica znaków kryptogramu i odpowiadających im znaków tekstu jawnego

Z tablicy, pokazanej w tabeli 6, otrzymuje się natychmiast tablicę, którą należy stosować przy deszyfrowaniu (tabela 7).

Można teraz zademonstrować proces tworzenia kryptogramu szyfru podstawieniowego. Jeśli tekstem jawnym będzie blok, pokazany w tabeli 2, to po zastosowaniu tablicy z tabeli 6 i wykonaniu 48 operacji podstawienia znaków otrzyma się kryptogram jak w tabeli 8:

X	U	Q	I	,	:	E	
W	Z	Y	E	I	D	Z	D
X	R	B	D	J	K	T	
X	C	V	B	D	Q	;	T
X	:	E	B	D	:	X	
X	Y	E	D	Z	;	;	M

Tabela 8. Blok kryptogramu szyfru podstawieniowego, odpowiadający blokowi tekstu jawnego

Otrzymany kryptogram można jednak bardzo łatwo rozszyfrować, bo występujące w kryptogramie znaki występują w takiej samej kolejności, jak odpowiadające im znaki w tekście jawnym. Kryptoanalitycy łamią wiec szyfry podstawieniowe bardzo szybko.

7. Przykład szyfru produktowego

Szyfrem produktowym jest szyfr, który stosuje po kolei dwa algorytmy szyfrujące: najpierw szyfruje się tekst jawny za pomocą algorytmu pierwszego, a następnie otrzymany kryptogram szyfruje się, stosując algorytm drugi. W efekcie otrzymuje się kryptogram wynikowy szyfru produktowego.

X	I	D	X	D	:	U	X
Z		C	V	X	Q	I	Y
E	,	D	B	:	D	E	Z
D	X		Q	Q	;	;	T
B	X	D	J	W	;	K	:
E	T	Z		Y	E	B	M

Tabela 9. Blok kryptogramu szyfru produktowego, otrzymany przez zaszyfrowanie bloku tekstu jawnego, pokazanego w tabeli 2, najpierw za pomocą szyfru przestawieniowego, a następnie za pomocą szyfru podstawieniowego

Można tę zasadę wyjaśnić, szyfrując tekst jawny z tabeli 2 najpierw za pomocą opisanego wyżej szyfru podstawieniowego, a następnie otrzymany w tym etapie kryptogram (tabela 3) traktuje się jako wejście dla algorytmu szyfrowania szyfrem podstawieniowym. Otrzymany w efekcie tych działań kryptogram pokazano w tabeli 9. Można z łatwością sprawdzić, ze deszyfrowanie kryptogramu z tabeli 9 należy przeprowadzić w kolejności odwrotnej do procesu szyfrowania: najpierw zastosować trzeba algorytm deszyfrowania dla szyfru podstawieniowego, a następnie algorytm deszyfrowania szyfru przestawieniowego.

Jakkolwiek składnikami opisanego tu algorytmu szyfrowania produktowego są dwa słabe szyfry, to złamanie przedstawionego szyfru produktowego, nawet dla zaawansowanego kryptoanalityka, nie będzie zadaniem trywialnym.

Właśnie takie naprzemienne stosowanie szyfrów przestawieniowych i podstawieniowych, prowadzi, zgodnie ze wskazówkami Shannona, do syntezy odpornych do złamania szyfrów. Te nieco zmodyfikowaną zasadę stosuje wiele praktycznie używanych systemów kryptograficznych dla szyfrowania blokowego, operujących na alfabecie dwuelementowym, m. in. algorytmy DES i IDEA.