Streszczenie

Informacje ukrywano przed innymi osobami jeszcze przed rozpowszechnieniem się pisma. Obecnie szyfrowanie informacji ma zarówno zastosowania militarne, jak i cywilne, np. przy przesyłaniu w sieci danych, których właściwego znaczenia nie powinien poznać nikt poza nadawcą i odbiorcą, nawet jeśli wiadomość wpadnie w niepożądane ręce. Wykład jest wprowadzeniem do kryptografii, ze szczególnym uwzględnieniem kryptografii komputerowej. Na początku zostaną omówione ciekawe metody szyfrowania, którymi posługiwano się przed erą komputerów, a które są wykorzystywane jeszcze dzisiaj. W drugiej zaś części wykładu zostanie zaprezentowany algorytm szyfrowania z kluczem publicznym, który jest wykorzystywany na przykład przy wysyłaniu rozwiązań zadań z Olimpiady Informatycznej. Za początek ery kryptografii komputerowej uważa się osiągnięcia zespołu polskich kryptoanalityków pod kierunkiem Mariana Rejewskiego, którzy złamali szyfr niemieckiej maszyny Enigma. Wspomniane będzie o tym krótko w czasie wykładu.

Fragmenty tego opracowania pochodzą z podręcznika [3].

Spis treści

1 WPROWADZENIE

Człowiek od początku swoich dziejów porozumiewał się z innym człowiekiem. Z czasem, wynalazł różne sposoby komunikowania się na odległość. Później, rodzące się społeczności, państwa i ich wodzowie potrzebowali sprawnych systemów łączności, by skutecznie władać swoimi, często rozległymi posiadłościami. Sprzeczne interesy różnych władców i dowódców powodowały, że zaczęli oni strzec wysyłanych przez siebie wiadomości przed przechwyceniem ich przez inne osoby. Tak narodziła się potrzeba **szyfrowania**, czyli utajniania treści przesyłanych wiadomości, i doprowadziło to do powstania **kryptografii**, dziedziny wiedzy, a niektórzy twierdzą, że także dziedziny sztuki, zajmującej się pierwotnie szyfrowaniem.

Wielokrotnie w historii ludzkości szyfrowanie miało istotny wpływ na bieg wydarzeń. Najbardziej spektakularnym przykładem jest chyba historia rozpracowania niemieckiej maszyny szyfrującej Enigma, dzięki czemu – jak utrzymują historycy – II wojna światowa zakończyła się o 2-3 lata wcześniej. Dużą w tym rolę odegrali Polacy (patrz rozdz. 7).

Sposoby utajniania wiadomości i łamania zabezpieczeń stanowią ze swej natury wiedzę tajemną – wiele faktów z tej dziedziny jest długo utrzymywanych w tajemnicy lub wcale nieujawnianych.

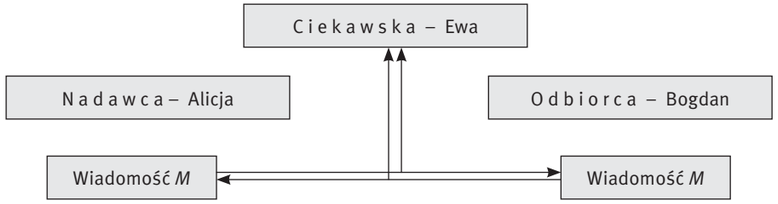
> „Ofiarami takiego stanu rzeczy stali się dwaj najbardziej zasłużeni dla informatyki Brytyjczycy, Charles Babbage i Alan Turing. Około 1854 roku, Babbage opracował sposób łamania szyfru Vigenère'a, który od połowy XVI wieku był powszechnie stosowany w łączności dyplomatycznej i wojskowej i uchodził za szyfr nie do złamania. O swoich pracach w tej dziedzinie Babbage nie wspomina nawet w swojej biografii – dowiedziano się o tym dopiero z jego korespondencji odczytanej sto lat później. Całe zasługi na tym polu przypisuje się Fridrichowi W. Kasiskiemu, który złamał szyfr Vigenère'a w 1863 roku. Historycy tłumaczą to dwojako: w zwyczaju Babbage'a było niekończenie prac, nie opublikował on więc swoich wyników, ale niektórzy twierdzą, że jego osiągnięcie utajnił wywiad brytyjski w związku z trwającą Wojną Krymską. Z kolei Alan Turing, w czasie II Wojny Światowej uczestniczył m.in. w pracach wywiadu brytyjskiego nad łamaniem szyfrów niemieckich, tworzonych za pomocą maszyny Enigma, bazował przy tym na osiągnięciach Polaków. Prawo w Wielkiej Brytanii chroni jednak informacje wywiadowcze przez przynajmniej 30 lat – świat dowiedział się więc o jego osiągnięciach dopiero w 1974 roku, gdy nie żył on już od ponad 20 lat. Okazało się wtedy również, że już w 1943 roku Brytyjczycy skonstruowali komputer Colossus oparty na przekaźnikach, by łamać najbardziej tajne szyfrogramy Hitlera. Colosuss powstał zatem dwa lata wcześniej niż maszyna ENIAC, uważana powszechnie za pierwszy komputer elektroniczny" [8].

Wraz z ekspansją komputerów i Internetu, dane i informacje są coraz powszechniej przechowywane i przekazywane w postaci elektronicznej. By nie miały do nich dostępu nieupoważnione osoby, szyfrowane są zarówno dane przechowywane w komputerach, jak i tym bardziej dane i informacje przekazywane drogą elektroniczną, np. poczta elektroniczna, rozmowy telefoniczne (rozmowy prowadzone za pomocą komórek są kierowane tylko do wybranych odbiorców, inni nie mogą ich usłyszeć) czy operacje bankowe wykonywane z automatów bankowych lub w formie płatności elektronicznych.

Obecnie kryptografia obejmuje znacznie szerszy zakres zagadnień niż tradycyjnie rozumiane szyfrowanie wiadomości i zajmuje się również identyfikacją użytkowników sieci oraz podpisem elektronicznym i znajduje zastosowanie przy kontroli dostępu użytkowników do stron i serwisów internetowych, do banków, w realizacji płatności elektronicznych przy zakupach internetowych, przy zabezpieczaniu prawa do własności intelektualnych i w wielu jeszcze innych obszarach komunikacji z wykorzystaniem sieci komputerowych, a ogólniej – technologii informacyjno-komunikacyjnych.

1.1 KOMUNIKACJA, SZYFRY I ICH RODZAJE

Ogólny schemat sytuacji, w której pojawia się konieczność szyfrowania, jest przedstawiony na rysunku 1. Alicja komunikuje się z Bogdanem lub Bogdan z Alicją, ale rozmowa może być podsłuchiwana przez ciekawską Ewę. Dalej będziemy używać tych imion w opisie sytuacji komunikacyjnej. Poczyńmy także uproszczenie, że interesować nas będzie bezpieczne przesyłanie wiadomości od Alicji do Bogdana. Przesyłanie wiadomości w drugą stronę odbywa się identycznie.



Rysunek 1.
Komunikowanie się Alicji z Bogdanem i ciekawska Ewa w pobliżu

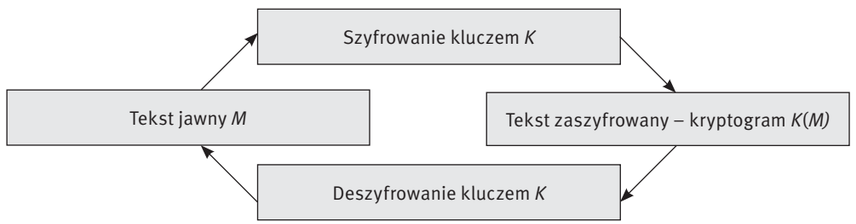
Aby Ewa nie mogła poznać wiadomości przesyłanych przez Alicję do Bogdana, muszą oni w jakiś sposób utajniać wysyłane wiadomości, czyli je **szyfrować**. **Szyfr**, to ustalony sposób utajniania (czyli **szyfrowania**) wiadomości, który na ogół polega na zastępowaniu tekstu wiadomości innym tekstem, z którego trudno domyśleć się znaczenia tekstu oryginalnego. Wiadomość, którą utajniamy określa się mianem **tekstu jawnego**, a jego zaszyfrowaną wersję – **kryptogramem**. Odczytywanie tekstu jawnego z szyfrogramu nazywa się **deszyfrowaniem**.

Tekst można szyfrować na dwa sposoby:

- przestawiając tylko znaki – jest to szyfrowanie **przez przestawianie**, czyli tzw. **szyfr przestawieniowy**;
- albo zastępując znaki tekstu innymi znakami – jest to szyfrowanie **przez podstawianie**, czyli tzw. **szyfr podstawieniowy**.

Najpowszechniej są stosowane **metody podstawieniowe**, które polegają na zamianie znaków (liter) innymi znakami (literami), czyli na podstawianiu za nie innych liter. A zatem, litery tworzące wiadomość nie zmieniają swojego miejsca w wiadomości, tylko zmieniają swoje znaczenie.

Istniejące i stosowane w praktyce metody szyfrowania są na ogół dobrze znane – na czym więc polega utajnianie wiadomości? Nieujawnianym elementem metod szyfrowania jest ich **klucz**, który służy do „otwarcia" utajnionej wiadomości. Na rysunku 2 przedstawiono bardzo ogólny schemat szyfrowania z kluczem *K*. Zauważmy, że ten sam klucz *K* jest wykorzystany do deszyfrowania wiadomości – jest to tak zwana **symetryczna metoda szyfrowania**. W dalszej części omówimy również **metodę asymetryczną**, w której inny klucz jest użyty do szyfrowania, a inny do deszyfrowania (patrz rozdz. 8).



Rysunek 2.
Schemat metody szyfrowania i deszyfrowania

1.2 KODOWANIE JAKO SZYFROWANIE

Jedną z odmian szyfrowania metodą podstawiania jest **kodowanie** wiadomości. Kodowanie polega na ogół na zastępowaniu znaków tworzących oryginalną wiadomość innymi znakami lub zbiorami znaków według jednoznacznego przepisu, danego zazwyczaj w postaci **książki kodowej**, którą można uznać za klucz w tej metodzie. Przykładami takiej „książki" jest kod ASCII, który służy do zamiany znaków na bajty, oraz alfabet Morse'a, w którym znaki są zastępowane ciągami kropek i kresek. Kodowanie może służyć ukrywaniu znaczenia wiadomości, a w kryptografii komputerowej przedstawianie wiadomości w kodzie ASCII jest pierwszym etapem szyfrowania wiadomości.

W czasie II wojny światowej w wojsku amerykańskim był stosowany kod Nawahów, w którym każdej literze alfabetu odpowiadało słowo z języka tego plemienia Indian. W oddziałach zostali również zatrudnieni łącznościowcy, wywodzący się tego plemienia, którzy jako jedyni byli w stanie szyfrować i deszyfrować wiadomości przesyłane za pomocą słów z ich języka.

2 POCZĄTKI KRYPTOGRAFII

Zanim przedstawimy wybrane aspekty współczesnej kryptografii, opiszemy kilka sposobów utajniania wiadomości, stosowanych w przeszłości. Osobom zainteresowanym szerszym omówieniem historii kryptografii polecamy bardzo ciekawie napisaną książkę Simona Singha, *Księga szyfrów* [8].

2.1 STEGANOGRAFIA

Terminem **steganografia** określa się ukrywanie przekazywanych wiadomości. Pierwsze wzmianki na ten temat znajdujemy w historii zmagań między Grecją a Persją w V w. p.n.e., opisanych przez sławnego historyka starożytności, Herodota. Steganografia przybierała różne formy, np. wiadomość zapisywano na cienkim jedwabiu, robiono z niego kulkę i po oblaniu jej woskiem goniec ją połykał. Inną formą steganografii jest stosowanie sympatycznego atramentu – tekst nim zapisany znika, ale pojawia się na ogół po podgrzaniu papieru lub potraktowaniu go specjalną substancją. Sławne stało się zapisanie wiadomości na ogolonej głowie wojownika i wysłanie go z tą wiadomością, gdy odrosły mu włosy.

Steganografia ma podstawową wadę – przechwycenie wiadomości jest równoznaczne z jej ujawnieniem. Od pojawienia się łączności drogą bezprzewodową za pomocą fal radiowych zaczęła mieć marginalne znaczenie w systemach łączności. Obecnie czasem odżywana dzięki możliwościom miniaturyzacji – możliwe jest ukrycie nawet znacznej wiadomości, dodatkowo wcześniej zaszyfrowanej, np. w... kropce, umieszczonej w innym tekście.

Polecamy bardzo ciekawy artykuł na temat steganografii, zamieszczony w czasopiśmie „Wiedza i Życie" 6/2002, s. 38-42, oraz strony w Internecie poświęcone ukrywaniu wiadomości. Obecnie wraca popularność steganografii wykorzystywanej do ukrywania wiadomości, np. znaków tekstu, w „wolnych miejscach" innych plików lub fragmentów wiadomości w postaci niewidocznego tła dokumentu.

W praktyce, dla zwiększenia bezpieczeństwa stosuje się steganografię w połączeniu z kryptografią, czyli ukrywa się przesyłanie utajnionych wiadomości.

2.2 SZYFR CEZARA

Jednym z najstarszych znanych szyfrów podstawieniowych jest **szyfr Cezara**, pochodzący z I w. p.n.e. Polega on na zastąpieniu każdej litery tekstu jawnego literą położoną w alfabecie o trzy miejsca dalej. Liczba 3 jest więc **kluczem** dla klasycznego szyfru Cezara. Kluczem w tym sposobie szyfrowania może być jednak dowolna liczba między 1 a długością alfabetu, o którą przesuwamy każdą literę alfabetu.

W przypadku szyfrów podstawieniowych można mówić o **alfabecie jawnym**, w którym są zapisywane wiadomości, i o **alfabecie szyfrowym**, w którym są zapisywane utajniane wiadomości. Dla ułatwienia umieszczamy te alfabety jeden pod drugim i tekst jawny będziemy pisać małymi literami, a zaszyfrowany – wielkimi. A zatem w przypadku szyfru Cezara mamy (oba alfabety zawierają również polskie litery):

Alfabet jawny **a ą b c ć d e ę f g h i j k l ł m n ń o ó p q r s ś t u v w x y z ż ź**
Alfabet szyfrowy **C Ć D E Ę F G H I J K L Ł M N Ń O Ó P Q R S Ś T U V W X Y Z Ż Ź A Ą B**

Polecamy zaszyfrowanie powiedzenia Cezara: **veni! vidi! vici!** stosując jego szyfr[19]. Jakie przysłowie łacińskie zaszyfrowano w ten sposób w następującym kryptogramie:

TGSGWLWLQ GUW OCWGT UWXFLQTXO – SQZWCTĄCÓLG ŁGUW OCWMĆ XEAĆEŹEK ULH

Wiedząc, że szyfrowano tekst z przesunięciem 3, łatwo można odczytać utajnioną wiadomość, gdyż jest to operacja odwrotna – litery szyfrogramu znajdujemy w alfabecie szyfrowym, a ich odpowiedniki w alfabecie jawnym. Szyfr Cezara jest więc szyfrowaniem z **kluczem symetrycznym**, gdyż znajomość klucza nadawcy wystarcza do określenia klucza odbiorcy. Niestety, tak zaszyfrowany tekst może rozszyfrować osoba nieuprawniona po przechwyceniu zaszyfrowanej wiadomości wiedząc tylko, że nadawca zastosował szyfr Cezara.

Klasyczny szyfr Cezara można uogólnić, dopuszczając jako alfabet szyfrowy przesunięcie alfabetu jawnego o dowolną liczbę znaków – mamy więc 34 możliwe alfabety w przypadku alfabetu języka polskiego. To jednak nadal niewielkie utrudnienie i nawet mało wprawiony **kryptolog**, czyli „łamacz szyfrów", po przechwyceniu kryptogramu zaszyfrowanego uogólnionym szyfrem Cezara łatwo określi, jaki był alfabet szyfrowy. Spróbuj i Ty swoich sił i rozszyfruj przysłowie łacińskie, które zostało ukryte w następującym kryptogramie:

RSBCXASK ESCKO VKQSBCAK OBC

otrzymanym za pomocą uogólnionego szyfru Cezara dla alfabetu łacińskiego. *Uwaga*. Alfabet łaciński nie zawiera polskich liter, nie zawiera również litery J, a więc składa się z 25 liter.

2.3 SZYFR PLAYFAIRA

Przedstawimy tutaj szyfr, zwany **szyfrem Playfaira**, który w połowie XIX wieku wymyślił Charles Wheatstone, jeden z pionierów telegrafu, a spopularyzował Lyon Playfair. Jest to przykład szyfru podstawieniowego, w którym pary różnych liter tekstu jawnego są zastępowane inną parą liter. Aby móc stosować ten szyfr, nadawca i odbiorca muszą jedynie wybrać słowo kluczowe. Przypuśćmy, że jest nim słowo **WYPAD**. Na podstawie słowa kluczowego jest tworzona tabliczka – patrz tabela 1 – złożona z 25 (5 na 5) pól, służąca do szyfrowania i deszyfrowania wiadomości. Umieszczamy w niej najpierw słowo kluczowe, a następnie pozostałe litery alfabetu łacińskiego (**I** oraz **J** w jednym polu; i w tekstach zaniedbujemy znaki diakrytyczne w polskich literach).

Tabela 1.
Przykładowa tabliczka-klucz dla metody szyfrowania Playfaira

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

Tekst jawny przed zaszyfrowaniem dzielimy na pary różnych liter – takie same litery przedzielamy np. literą **x**, i ostatnią pojedynczą literę również uzupełniamy do pary literą **x**.

Tekst jawny: **do zobaczenia o szóstej** przyjmuje więc przed szyfrowaniem postać: **do-zo-ba-cz-en-ia--os-zo-st-ej**. A oto reguły zastępowania par liter:

- jeśli obie litery znajdują się w tym samy wierszu, np. **os**, to zastępujemy je sąsiednimi literami z prawej strony; jeśli jedna z liter znajduje się na końcu wiersza, w naszym przypadku jest to **s**, to zastępujemy ją

[19] Dla uproszczenia, znaki inne niż litery (np. cyfry, znaki interpunkcyjne) pomijamy w kryptogramach, z wyjątkiem odstępów w dłuższych tekstach.

pierwszą literą wiersza; a zatem w tym przypadku **os** zastępujemy parą **QN**;
- jeśli obie litery znajdują się w tej samej kolumnie, to zastępujemy je literami leżącymi bezpośrednio poniżej; jeśli jedna z liter znajduje się na końcu kolumny, to zastępujemy ją pierwszą literą kolumny;
- jeśli obie litery nie znajdują się ani w tym samym wierszu, ani w tej samej kolumnie, to zastępujemy je parą według schematu pokazanego powyżej, w odniesieniu do pary **do** – zastępujemy ją parą **YS**; w obu przypadkach, polega to na wzięciu litery znajdującej się w tym samym wierszu i w kolumnie, w której znajduje się druga litera.

Ostatecznie, tekst **do zobaczenia o szóstej** zostanie zaszyfrowany jako:

YS-US-FW-GU-BQ-LY-QN-US-NZ-CK

A jaki tekst jawny został zaszyfrowany dla tego samego klucza w następującym kryptogramie:

KCRMKHDRUGNBKMBVNYCOGDRWQDYFFQVDQGLZ

3 SZYFROWANIE PRZEZ PRZESTAWIANIE

Na chwilę tylko zatrzymamy się przy szyfrowaniu przez przestawianie, gdyż jest to mało popularny sposób utajniania wiadomości, chociaż przestawienie liter w tekście wiadomości jest jednym z najprostszych sposobów zmiany znaczenia wiadomości. Przestawiając litery w słowie możemy otrzymać inne słowo, zwane jego **anagramem**. A zatem taki sposób szyfrowania dłuższego tekstu jest uogólnieniem anagramu.

Znajdź dwa anagramy „słowa" **AGLMORTY**[20], które wielokrotnie pojawiają się na zajęciach z tego działu informatyki w projekcie Informatyka +. Czy nie dziwi Cię, że są one swoimi anagramami?

Jeśli w danym tekście możemy dowolnie przestawiać litery, to otrzymujemy olbrzymią liczbę możliwych kryptogramów. Przypuśćmy, że chcemy zaszyfrować słowo **szyfrowanie** przestawiając tylko litery tego słowa. Wszystkie możliwe przestawienia liter tego słowa można otrzymać budując słowa litera po literze. W przypadku słowa **szyfrowanie** mamy do dyspozycji litery ze zbioru {**a, e, f, i, n, o, r, s, w, y, z**}. Aby utworzyć słowo z liter tego zbioru (słowo to nie musi mieć znaczenia), najpierw wybieramy pierwszą literę (spośród 11), następnie określamy drugą literę, wybierając ją spośród 10 pozostałych, dla trzeciej litery w słowie mamy 9 możliwości itd. Zatem wszystkich możliwości jest: 11*10*9*...*2*1 = 39916800. Ten iloczyn oznaczamy przez 11! – jest to funkcja zwaną **silnią**. Ogólnie mamy:

$$n! = 1*2*3* \ldots*(n-1)*n.$$

Tabela 2.
Wartości funkcji *n*! dla przykładowych *n*

n	*n*!
5	120
15	$1{,}3*10^{12}$
25	$1{,}5*10^{25}$
35	$1{,}0*10^{40}$
50	$3{,}0*10^{64}$

W tabeli 2 są podane wartości funkcji silnia dla kilku wartości *n*. Dla porównania, liczba protonów we wszechświecie jest szacowana na 10^{79}, a najszybsze obecnie komputery są w stanie wykonać w ciągu sekundy około 10^{15} operacji, a więc ta funkcja przyjmuje bardzo duże wartości i szybko rośnie ze wzrostem *n*.

[20] Zgodnie z przyjętą umową, wiadomości utajnione zapisujemy wielkimi literami, a wiadomości jawne – małymi literami.

Aby odczytać kryptogram, będący anagramem tekstu jawnego, należałoby wziąć pod uwagę wszystkie możliwe przestawienia liter. Jak pokazują liczby w tabeli 2, nawet dla tekstów o niewielkiej długości, bez pewnych reguł przestawiania będzie to szyfr tak samo trudny dla osoby przechwytującej kryptogram, jak i dla odbiorcy zaszyfrowanej wiadomości. Jedną z reguł tworzenia anagramu może być np. przestawianie liter na ustalonej pozycji.

Poniższy kryptogram został otrzymany za pomocą przestawienia każdych dwóch kolejnych liter w tekście jawnym poczynając do drugiej litery (zaniedbano odstępy między słowami).

WAHITILEKODNIBGSEITNSTOIHGN

Odczytaj tekst jawny (jest po angielsku). Czy pamiętasz, kto wypowiedział te słowa?

Przestawienia mogą dotyczyć bitów w dwójkowej reprezentacji tekstu jawnego. Można zaproponować na przykład następujący algorytm szyfrowania:

1. Zapisz wiadomość w kodzie ASCII.
2. Przestaw ze sobą co ósmą parę bitów poczynając od pary na pozycjach 2 i 3, a więc zamieniając ze sobą miejscami bity 2 i 3, 10 i 11, 18 i 19 itd. – w tym kroku definiuje się klucz tego sposobu szyfrowania.
3. Powróć, stosując kod ASCII, do wiadomości w postaci ciągu znaków.

Metoda płotu

Bardzo prosty szyfr przestawieniowy otrzymujemy stosując tzw. **metodę płotu**. W tej metodzie, najpierw kolejne litery tekstu jawnego są zapisywane na zmianę w dwóch rzędach, a następnie za kryptogram przyjmuje się ciąg kolejnych liter z pierwszego rzędu, po którym następuje ciąg kolejnych liter z drugiego rzędu. Na przykład, metoda płotu zastosowana do słowa **szyfrowanie** daje następujący kryptogram:

s y r w n e
z f o a i
SYRWNEZFOAI

„Płot” w tej metodzie szyfrowania może się składać z więcej niż dwóch rzędów, np. kryptogram powyższej wiadomości jawnej, otrzymany z użyciem trzech rzędów w „płocie”, ma postać: **SFWIZRAEYON**,

Dla dociekliwych i wytrwałych uczniów mamy następujące zadanie: Wiadomo, że poniższy kryptogram (powiedzenie premiera brytyjskiego Winstona Churchilla do szefa Secret Intelligence Service, 1941):

PIIŚIYOLSŁWETSIOŁŻZŁLŻEĄEESAŁ
ELEERCKNNDMŹZKOEMBOOOAIZŻMTDNCCYBTMAEIEIOON

powstał z tekstu jawnego za pomocą metody płotu. Nie wiadomo jednak, z ilu rzędów składał się „płot”. Rozszyfruj tę wiadomość. Podpowiemy tylko: zauważ, że umieszczając tekst jawny w kilku rzędach płotu, każdy rząd płotu zawiera taką samą liczbę liter, z wyjątkiem kilku ostatnich rzędów, krótszych o jeden znak.

4 SZYFROWANIE Z ALFABETEM SZYFROWYM

Wracamy tutaj do szyfrowania z alfabetem szyfrowym – taką metodą jest opisany wcześniej szyfr Cezara. Za alfabet szyfrowy, zamiast przesuwania alfabetu o ustaloną liczbę liter, można wziąć jakiekolwiek uporządkowanie liter alfabetu. Takich uporządkowań jest jednak 35!, co jest olbrzymią liczbą. Niepowołana osoba ma więc małe szanse natrafić przypadkowo na wybrany z tej liczby alfabet szyfrowy. Jednak wybór losowego uporządkowania liter w alfabecie szyfrowym ma wadę. Alfabet ten trzeba w jakiś tajny sposób przekazać osobie, która ma odczytywać tworzone nim kryptogramy. Dlatego zamiast losowego wyboru alfabetu szyfrowego stosuje się różne rodzaje kluczy, które precyzyjnie określają alfabet szyfrowy.

Jednym ze sposobów określania alfabetu szyfrowego jest podanie **słowa** (lub **powiedzenia**) **kluczowego**, na podstawie którego tworzy się alfabet szyfrowy. Dla przykładu użyjmy **Cappadocia** jako słowo kluczowe do utworzenia alfabetu szyfrowego. Najpierw umieszczamy słowo kluczowe na początku alfabetu szyfrowego i usuwamy z niego powtarzające się litery – w naszym przykładzie pozostaje więc **CAPDOI**. Następnie, dopisujemy pozostałe litery w porządku alfabetycznym, zaczynając od litery następnej po ostatniej literze w pozostałej części słowa kluczowego (w przykładzie od litery **J**) i kontynuując od początku alfabetu. Otrzymujemy:

Alfabet jawny	**a ą b c ć d e ę f g h i j k l ł m n ń o ó p q r s ś t u v w x y z ż ź**
Alfabet szyfrowy	**C A P D O I J K L Ł M N Ń Ó Q R S Ś T U V W X Y Z Ż Ź Ą B Ć E Ę F G H**

Spróbuj odczytać, jaką wiadomość zaszyfrowano w poniższym tekście, posługując się szyfrem z kluczem **Cappadocia** (dla ułatwienia pozostawiliśmy w tekście odstępy i znaki interpunkcyjne):

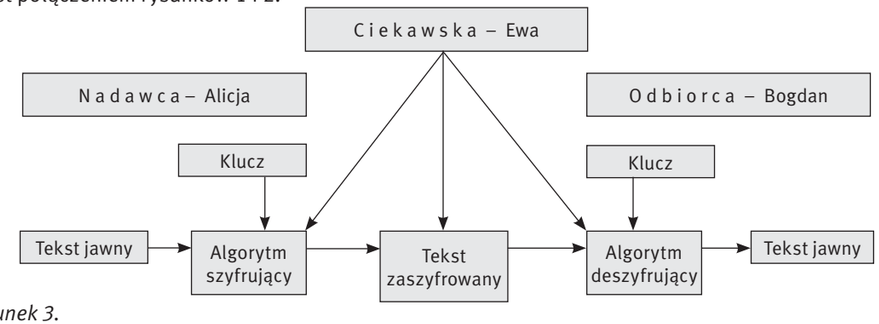
ÓCWCIUDŃC ŹU WYFJWNKÓŚC ÓYCNŚC ŚC ĆĘGĘŚNJ CŚCŹUQNŃZÓNJŃ Ć ŻYUIÓUĆJŃ
ŹĄYDŃN, UPLNŹĄŃADC Ć LCŚŹCZŹĘDFŚJ LUYSĘ ZÓCQŚJ, ĆĘIYAGUŚJ Ć ZÓCRCDM
ÓUŻDNURĘ N ÓQCZFŹUYĘ, IŃSĘ SNJZFÓCQŚJ, DCRJ WUIFNJSŚE SNCZŹC. ŹCS WUĆZŹCR
ŹJŚ LYCŁSJŚŹ WUIYKDFŚNÓC U ZFĘLYUĆCŚNĄ.

W czasie II wojny światowej w polskim ruchu oporu stosowano alfabety szyfrowe, określane na podstawie ustalonych fragmentów wybranych książek. Na przykład za klucz można przyjąć siedem pierwszych słów z pierwszego akapitu na 267 stronie *Potopu*, wydanego w 1923 roku we Lwowie (obie komunikujące się strony musiały posiadać dokładnie tę samą książkę), a w naszym przypadku może to być na przykład pierwsze zdanie ze streszczenia tych materiałów – podaj, jaki to będzie klucz.

Tworzenie alfabetów szyfrowych na podstawie słów kluczowych znacznie utrudnia deszyfrację kryptogramów, gdyż możliwych jest bardzo wiele słów (a ogólniej tekstów) kluczowych.

5 SCHEMAT KOMUNIKACJI Z SZYFROWANIEM

Podsumujemy tutaj krótko to, co dotychczas powiedzieliśmy o utajnianiu przesyłanych wiadomości. Rysunek 3 jest połączeniem rysunków 1 i 2.



Rysunek 3.
Schemat utajnionego przekazywania wiadomości

Przykładowe metody szyfrowania i deszyfrowania wiadomości, przedstawione w poprzednich punktach, można opisać schematycznie jak na rysunku 3. W tych metodach można wyróżnić następujące elementy:

- **algorytm szyfrowania**, który służy do zamiany tekstu jawnego na tekst zaszyfrowany (kryptogram) i **algorytm deszyfrujący**, służący do odszyfrowania kryptogramu na tekst jawny;

- **klucz**, który jest w pewnym sensie parametrem algorytmu szyfrującego i deszyfrującego, niezbędną informacją, w jaki sposób mają działać algorytmy szyfrujący i deszyfrujący.

We współczesnej kryptografii przyjmuje się założenie, że algorytmy szyfrowania są znane, natomiast ukryty jest klucz do algorytmów szyfrowania i deszyfrowania, a zatem Ciekawska Ewa zna sposób utajniania wiadomości, które przesyłają między sobą Alicja i Bogdan, ale nie zna klucza, który jest używany przez nich w algorytmach szyfrujących i deszyfrujących. Jest to odzwierciedleniem podstawowego założenia, jakie przyjmuje się obecnie w kryptografii:

> bezpieczeństwo systemu szyfrowania jest oparte na kluczach, a nie na sposobie szyfrowania.

Przypomnijmy więc sobie, jakie były klucze w poznanych algorytmach szyfrowania:

- szyfr Cezara – 3 lub inna liczba, o którą przesuwa się alfabet jawny, by otrzymać alfabet szyfrowy;
- metoda płotu – liczba rzędów w płocie;
- szyfr Playfaira – słowo kluczowe, które służy do zbudowania tabliczki zamiany par liter;
- szyfr z jednym alfabetem szyfrującym, tzw. **szyfr monoalfabetyczny** – słowo kluczowe (lub tekst kluczowy), które służą do zbudowania alfabetu szyfrowego.

Zauważmy, że we wszystkich tych przypadkach ten sam klucz służy do szyfrowania wiadomości i do deszyfrowania kryptogramów. Stwarza to pewne problemy z przekazywaniem klucza, co musi być czynione z wielką ostrożnością, by klucz nie wpadł w ręce Ewy.

Podane przez nas przykłady użycia powyższych algorytmów pokazują, że znalezienie klucza w przypadku pierwszego i drugiego algorytmu nie jest specjalnie trudne, zwłaszcza jeśli dysponujemy komputerem. W dwóch pozostałych przypadkach nie jest to już jednak takie proste, gdyż istnieją nieograniczone możliwości doboru słów kluczowych. Jednak i te szyfry udało się złamać.

Łamaniem szyfrów, czyli odczytywaniem utajnionych wiadomości bez znajomości wszystkich elementów sposobu szyfrowania (np. bez znajomości klucza), zajmuje się **kryptoanaliza**. Kryptografia i kryptoanaliza to dwa główne działy **kryptologii**, nauki o utajnionej komunikacji. Kryptografia i kryptoanaliza rywalizują ze sobą od początków utajniania wiadomości – kryptografia tworzy coraz doskonalsze (bezpieczniejsze) szyfry, a kryptoanaliza dostarcza metod ich łamania. Jak pokazaliśmy, złamanie uogólnionego szyfru Cezara jest dość proste – należy jedynie ustalić, o ile pozycji przesuwają się wszystkie litery. Podobnie jest ze złamaniem szyfru płotowego – należy określić, jak wysoki jest płot, czyli ile ma rzędów. W następnym punkcie opowiemy o łamaniu szyfrów z jednym alfabetem szyfrowym i o pewnym uogólnieniu tej metody.

6 SZYFR POLIALFABETYCZNY

Wracamy tutaj do sposobu szyfrowania z alfabetem szyfrującym, gdyż interesuje nas, czy można złamać taki szyfr z alfabetem szyfrowania, utworzonym przez dowolne słowo kluczowe. Jeśli tak, to na pewno nie jest to metoda sprawdzająca wszystkie możliwe alfabety szyfrowania, gdyż jest ich zbyt dużo. Pierwszy zapis o skutecznym sposobie łamania szyfru z alfabetem szyfrowania pochodzi z IX wieku. Jego autorem był Al-Kindi, zwany „filozofem Arabów". Jako pierwszy wykorzystał on różnice w częstości występowania liter w tekstach i zastosował tzw. **metodę częstościową**. Wyjaśnimy krótko, na czym polega ta metoda.

W tabeli 3 są podane częstości występowania liter w tekstach w języku angielskim i w języku polskim. W tekstach angielskich zdecydowanie najczęściej występuje litera **e**, a następnie litery **t**, **a**, **o**, **i**. W tekstach w języku polskim natomiast żadna litera nie występuje tak często jak litera **e** w tekstach angielskich, a z częstością ponad 7% występują litery: **a**, **i**, **o**, **e**. Stąd można wywnioskować, że litera, która występuje najczęściej w kryptogramie w języku polskim powinna odpowiadać literze **a** w tekście jawnym. Oczywiście odnosi się to do nieco dłuższych tekstów, a więc w krótszych kryptogramach ta najczęściej występująca w nich litera może odpowiadać którejś z następnych co do częstości pojawiania się w tekstach liter. Ponadto, szukając par odpowiadających sobie liter uwzględnia się charakterystyczne dla danego języka połączenia dwóch lub więcej liter, np. w języku polskim dość często występują pary liter: **rz**, **sz**, **cz**, **ść**, a w języku angielskim – **qu**, **th**. Sposób łamania szyfru, czyli deszyfracji wiadomości, polegający na wykorzystaniu częstości występowania liter i ich kombinacji w tekście nazywa się **analizą częstości**.

Tabela 3.
Częstości występowania liter w tekstach w języku angielskim i polskim

Litery	Częstość występowania liter w tekstach (w %)	
	w języku angielskim	w języku polskim
A	8,1	8,71
B	1,5	1,29
C	2,8	3,91
D	4,3	3,45
E	12,6	7,64
F	2,2	0,38
G	2,0	1,42
H	6,1	1,22
I	7,0	8,48
J	0,2	2,38
K	0,8	3,10
L	4,0	2,09
M	2,4	2,64
N	6,7	5,60
O	7,5	7,90
P	1,9	3,01
Q	0,1	–
R	6,0	4,63
S	6,3	4,64
T	9,0	3,63
U	2,8	1,92
V	1,0	–
W	2,4	4,62
X	0,2	–
Y	2,0	3,88
Z	0,1	5,95

Jeśli policzymy częstości poszczególnych liter w kryptogramie podanym powyżej dla słowa kluczowego Cappadocia to się okaże, że odpowiedniki pięciu najczęściej występujących liter w tym kryptogramie plasują się na pierwszych (pod względem częstości) sześciu pozycjach w ostatniej kolumnie tabeli 3.

Posłużenie się analizą częstości przy deszyfrowaniu tekstu nie jest tylko prostym skorzystaniem z tabeli częstości liter, ale żmudną analizą możliwych przypadków, w której nierzadko trzeba postępować metodą prób i błędów, czyli próbować różnych przyporządkowań liter i nieraz się z nich wycofywać. Procesu deszyfracji, wykorzystującego analizę częstości, nie daje się prosto zautomatyzować za pomocą komputera. Jest to postępowanie w dużym stopniu interaktywne.

Po złamaniu monoalfabetycznego szyfru podstawieniowego, czyli z jednym alfabetem szyfrowym, nastąpił ruch kryptografów, którzy zaproponowali szyfr z wieloma alfabetami szyfrowymi – taki szyfr nazywa się **polialfabetycznym**. Uniemożliwia on użycie prostej analizy częstości, gdyż w tworzonych kryptogramach za daną literę może być podstawianych wiele różnych liter. Słowo kluczowe w tym przypadku służy do określenia alfabetów – są nimi alfabety Cezara wyznaczone przez kolejne litery słowa kluczowego. Dla przykładu, niech słowem kluczowym będzie **BRIDE** (dla utrudnienia kryptoanalizy, słowa kluczowe są wybierane z innych języków), wtedy mamy pięć alfabetów szyfrowych, zaczynające się od liter **B**, **R**, **I**, **D** i **E**.

Alfabet jawny	**a ą b c ć d e ę f g h i j k l ł m n ń o ó p q r s ś t u v w x y z ż ź**
Alfabety szyfrowe	**B C Ć D E Ę F G H I J K L Ł M N Ń O Ó P Q R S Ś T U V W X Y Z Ż Ź A Ą**
	R S Ś T U V W X Y Z Ż Ź A Ą B C Ć D E Ę F G H I J K L Ł M N Ń O Ó P Q
	I J K L Ł M N Ń O Ó P Q R S Ś T U V W X Y Z Ż Ź A Ą B C Ć D E Ę F G H
	D E Ę F G H I J K L Ł M N Ń O Ó P Q R S Ś T U V W X Y Z Ż Ź A Ą B C Ć
	E Ę F G H I J K L Ł M N Ń O Ó P Q R S Ś T U V W X Y Z Ż Ź A Ą B C Ć D

Zaszyfrujmy teraz wiadomość **panna w opałach!**: pierwszą literę jawnego tekstu szyfrujemy posługując się pierwszym alfabetem szyfrowym, drugą – drugim, trzecią – trzecim, czwartą – czwartym, piątą – piątym, szóstą – pierwszym itd. Aby nie zagubić się w liczeniu, którego kolejnego alfabetu użyć, dobrze jest umieścić powtarzające się słowo kluczowe nad tekstem do zaszyfrowania. W naszym przypadku mamy:

B R I D E B R I D E B R I
p a n n a w o p a ł a c h !

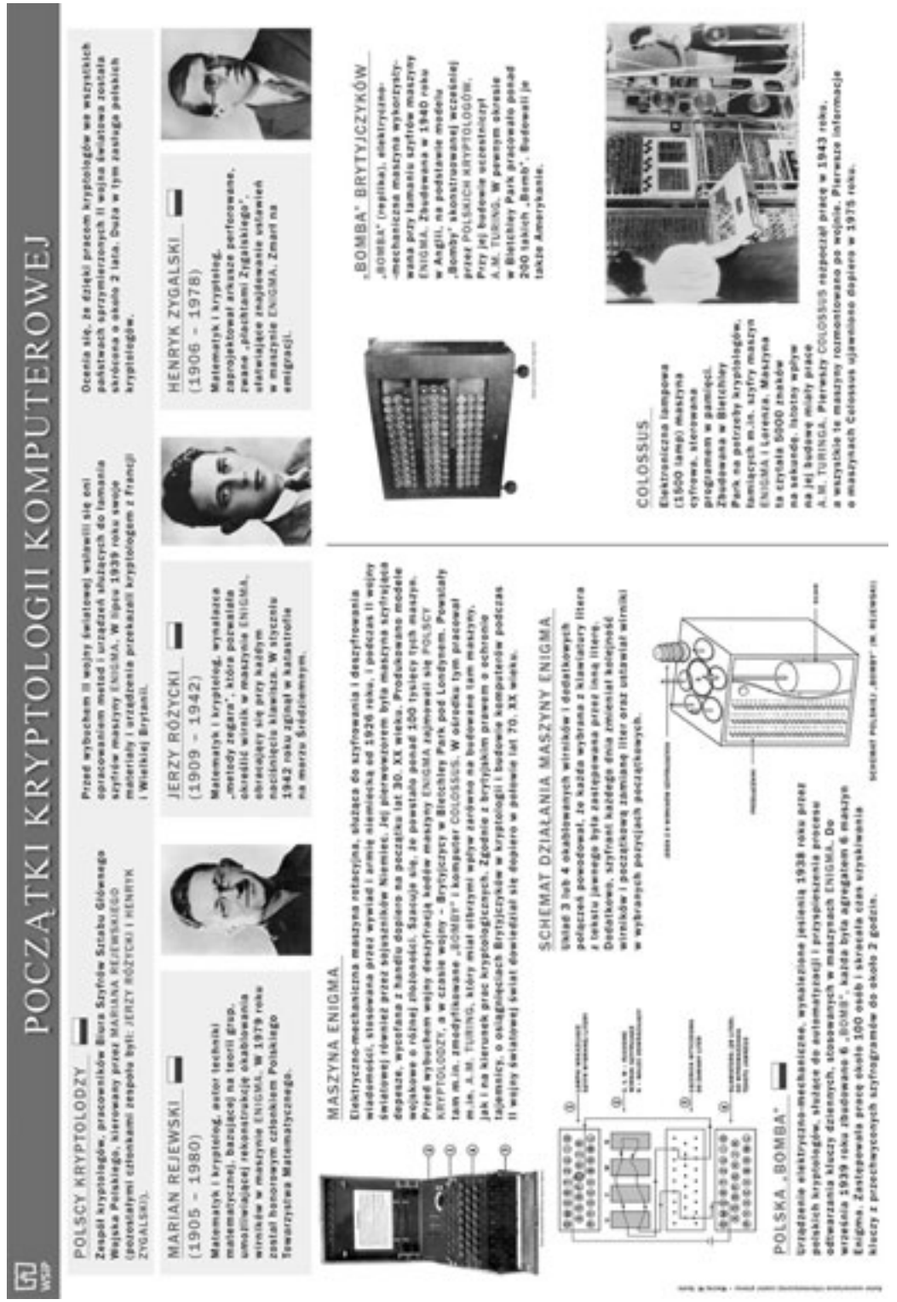
Otrzymujemy więc kryptogram: **RRVQE Y ĘZDPBTP!**, w którym każda z powtarzających się liter w tekście jawnym (**a** i **n**) jest za każdym razem szyfrowana przez inną literę.

Taki sposób szyfrowania został zaproponowany w XVI wieku przez dyplomatę francuskiego Blaise de Vigenère'a i nazywa się **szyfrem Vigenère'a**. Przekonaj się, że szyfrowanie tym sposobem nie przenosi częstości liter z tekstu jawnego na kryptogram. W tym celu zastosuj szyfr Vigenère'a z kluczem **BRIDE** do zaszyfrowania wiadomości jawnej, otrzymanej ze słowem kluczowym Cappadocia i oblicz częstości występowania liter w otrzymanym kryptogramie. Porównaj te częstości z częstościami liter w kryptogramie otrzymanym za pomocą jednego alfabetu z tym samy słowem kluczowym **BRIDE** oraz z częstościami ich odpowiedników w tekście jawnym (patrz tabela 3).

Dopiero w połowie XIX wieku kryptoanalitycy złamali szyfr Vigenère'a – jednym z ich był Charles Babbage, o czym wspominamy na początku tych materiałów. W odpowiedzi na złamanie szyfru Vigenère'a, kryptografowie zaproponowali jednoczesne szyfrowanie par liter – bardzo prosty szyfr tego rodzaju przedstawiliśmy już wcześniej, to **szyfr Playfaira**.

7 PRZEŁOM W KRYPTOGRAFII – ENIGMA

W latach dwudziestych pojawiła się pierwsza maszyna szyfrująca – **Enigma** – zastąpiła ona algorytm szyfrujący i deszyfrujący ze schematu na rysunku 3, a kluczem w tej maszynie było początkowe ustawienie jej elementów mechanicznych. Na początku maszynę tę można było nawet kupić na wolnym rynku. Szyfr Enigmy został złamany przez zespół polskich matematyków, którym kierował **Marian Rejewski**. Zbudowali oni maszynę deszyfrującą, zwaną **Bombą**, która składała się z wielu kopii Enigmy. Swoje osiągnięcia i pomysły wraz z tą maszyną Polacy przekazali jeszcze przed wybuchem II wojny światowej Brytyjczykom, którzy w Bletchley Park koło Londynu zapoczątkowali erę komputerowej kryptografii i kryptoanalizy. Na rysunku 4 przedstawiono krótką historię początków kryptografii komputerowej. Więcej na ten pasjonujący temat szyfrów maszyny Enigma można znaleźć w książkach: [2], [5], [6], [8].



Rysunek 4.
Początki kryptografii komputerowej w zarysie [ta i inne plansze dostępne są na stronie http://mmsyslo.pl/]

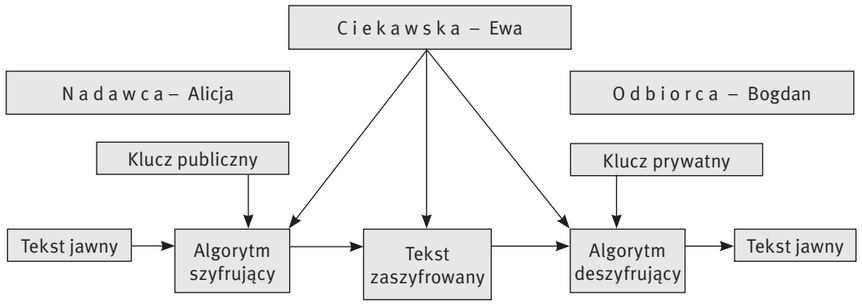
8 KRYPTOGRAFIA Z JAWNYM KLUCZEM

Pojawianie się coraz silniejszych komputerów powoduje realne zagrożenie dla przesyłania utajnionych wiadomości. Kryptoanalityk może skorzystać z mocy komputera, by sprawdzić bardzo dużą liczbę możliwych alfabetów i kluczy oraz prowadzić analizę kryptogramów metodą prób i błędów. Co więcej, z ekspansją komunikacji najpierw telefonicznej, a obecnie – internetowej wzrosła do olbrzymich rozmiarów liczba przesyłanych wiadomości. Państwa, instytucje, a także pojedynczy obywatele chcieliby mieć gwarancję, że system wymiany wiadomości może zapewnić im bezpieczeństwo i prywatność komunikacji.

W 1976 roku przyjęto w USA szyfr **Lucifer** jako standard komputerowego szyfrowania danych **DES** (ang. *Data Encryption Standard*). Liczba możliwych kluczy dla systemu DES jest gwarancją, że praktycznie jest to bezpieczny szyfr – powszechnie dostępne komputery są zbyt słabe, by go złamać. Pozostaje jednak nadal nierozwiązany tzw. **problem dystrybucji klucza** – w jaki sposób dostarczyć klucz odbiorcy utajnionej nim wiadomości. Problem ten dotyczył większości szyfrów w historii kryptografii. Na przykład Marian Rejewski osiągnął pierwsze sukcesy przy deszyfracji Enigmy, korzystając m.in. z faktu, że klucz do zaszyfrowanej wiadomości był dwa razy powtarzany na początku wiadomości, co okazało się słabą stroną kryptogramów Enigmy.

W połowie lat siedemdziesiątych pojawiła się sugestia, że wymiana klucza między komunikującymi się stronami być może nie jest konieczna. Tak zrodził się pomysł szyfru z jawnym kluczem, którego schemat jest przedstawiony na rysunku 5. Od schematu na rysunku 3 różni się tym, że zamiast jednego klucza dla nadawcy i odbiorcy mamy parę kluczy: **klucz publiczny**, zwanym **kluczem jawnym**, i **klucz prywatny**, zwany również **kluczem tajnym**. Jest to więc **szyfr asymetryczny**. Działanie odbiorcy i nadawcy utajnionych wiadomości w tym przypadku jest następujące:

1. Odbiorca wiadomości tworzy parę kluczy: publiczny i prywatny i ujawnia klucz publiczny, np. zamieszcza go na swojej stronie internetowej.
2. Jeśli nadawca chce wysłać zaszyfrowaną wiadomość do odbiorcy, to szyfruje ją jego kluczem publicznym i tak utworzony kryptogram może odczytać jedynie odbiorca posługując się kluczem prywatnym.



Rysunek 5.
Schemat przesyłania wiadomości w kryptografii z jawnym kluczem

Para kluczy – publiczny i prywatny – ma jeszcze tę własność, że znajomość klucza publicznego nie wystarcza nie tylko do odszyfrowania wiadomości nim zaszyfrowanej, ale również nie umożliwia utworzenia klucza prywatnego, który jest niezbędny do odszyfrowania wiadomości. Utworzenie takiej pary kluczy było możliwe dopiero w erze komputerów. Odbiorca może łatwo określić oba klucze z pary, ale nikt poza nim, przy obecnym stanie wiedzy i mocy komputerów nie jest w stanie odtworzyć klucza prywatnego na podstawie klucza publicznego. Ta trudność w odgadnięciu klucza prywatnego polega na tym, że jego znalezienie przez osobę postronną jest związane ze znalezieniem rozwiązania tak trudnego problemu, że żaden istniejący obecnie komputer nie jest w stanie w tym pomóc. Jeśli jednak pojawiłby się taki komputer, to łatwo można zwiększyć

trudność rozwiązania tego problemu przez niewielką modyfikację danych. Nieco szczegółów na ten temat zamieszczamy w rozdziale 10.

Najbardziej znany szyfr z kluczem publicznym opracowali w 1977 roku Ronald **Rivest**, Adi **Shamir** i Leonard **Adleman** z MIT i od inicjałów ich nazwisk pochodzi jego nazwa – **szyfr RSA**. Zilustrujemy jego działanie na niewielkim przykładzie komentując jednocześnie, jak wygląda jego pełna wersja (zob. również rys. 5). W poniższym opisie przyjmujemy, że Bogdan chce otrzymywać utajnione wiadomości od Alicji (patrz rys. 5 i 6), dlatego przygotowuje odpowiednie klucze dla siebie i dla Alicji.

Etap przygotowania kluczy

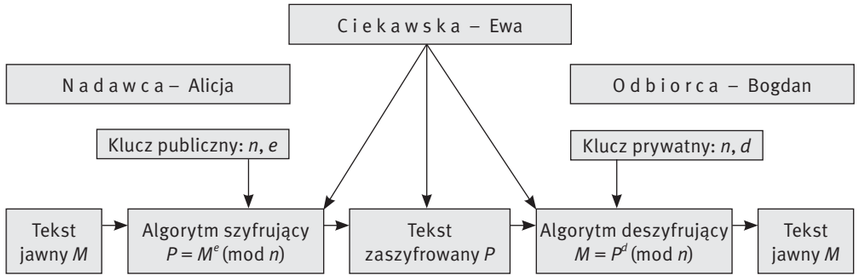
1. Bogdan wybiera dwie duże, możliwie bliskie sobie liczby pierwsze p i q. Liczby te powinny mieć przynajmniej po 200 cyfr. Liczby te Bogdan trzyma w tajemnicy.
   Przykład. Przyjmijmy $p = 11$ i $q = 13$.
2. Bogdan oblicza $n = p*q$ i wybiera dowolną liczbę naturalną e, która jest względnie pierwsza z liczbą $(p - 1)*(q - 1)$, czyli te dwie liczby nie mają wspólnych dzielników poza liczbą 1.
   Przykład. Mamy $n = 11*13 = 143$. Ponieważ $(p - 1)*(q - 1) = 10*12 = 120 = 2^3*3*5$, więc możemy przyjąć $e = 7$.
3. Bogdan znajduje liczbę naturalną d taką, że $e*d = 1 \pmod{(p - 1)*(q - 1)}$, czyli reszta z dzielenia $e*d$ przez $(p - 1)*(q - 1)$ jest równa 1. Liczbę d można znaleźć, posługując się algorytmem Euklidesa (patrz rozdz. 10). [Operacja mod k oznacza wzięcie reszty z dzielenia przez k.]
   Przykład. Mamy znaleźć liczbą naturalną d, spełniającą równość $e*d = 1 \pmod{(p - 1)*(q - 1)}$, czyli $7*d = 1 \pmod{120}$. Taką liczbą jest np. $d = 103$, gdyż $7*103 = 721$ i 721 podzielone przez 120 daje resztę 1.
4. Bogdan ogłasza parę liczb (n, e) jako swój klucz publiczny, np. zamieszcza go na swojej stronie w Internecie, a parę (n, d) zachowuje w tajemnicy jako klucz prywatny. Jednocześnie niszczy liczby p i q, by nie wpadły w niczyje ręce, co mogłoby umożliwić odtworzenie klucza prywatnego. Jeśli p i q są dostatecznie dużymi liczbami pierwszymi, to znajomość n i e nie wystarcza, by obliczyć wartość d posługując się nawet najpotężniejszymi dzisiaj komputerami (patrz rozdz. 10).

Szyfrowanie wiadomości

1. Jeśli Alicja chce wysłać do Bogdana jakąś wiadomość, to najpierw musi ją przedstawić w postaci liczby naturalnej M, nie większej niż n. Tekst można zamienić na liczbę posługując się kodem ASCII. Jeśli wiadomość jest zbyt długa, to należy ją szyfrować blokami odpowiedniej wielkości.
   Przykład. Jako wiadomość wybierzmy literę Q, która w kodzie ASCII ma kod 81. Przyjmujemy więc za wiadomość do wysłania $M = 81$.
2. Wiadomość, jaką Alicja wysyła do Bogdana, jest liczbą: $P = M^e \pmod{n}$. Obliczenie wartości P może wydawać bardzo złożone, w gruncie rzeczy można szybko wykonać przy wykorzystaniu jednej z szybkich metod potęgowania (patrz rozdz. 10) oraz własności operacji na resztach (zob. nasze przykładowe obliczenia).
   Przykład. Mamy obliczyć $P = 81^7 \pmod{143}$. Korzystając z rozkładu wykładnika na postać binarną mamy $81^7 = 81^1*81^2*81^4$. Z każdej z tych potęg obliczamy tylko resztę z dzielenia przez 143. Otrzymujemy stąd: $81^1*81^2*81^4 = 81*126*3 = 16 \pmod{143}$.
3. Alicja wysyła do Bogdana wiadomość $P = 16$.

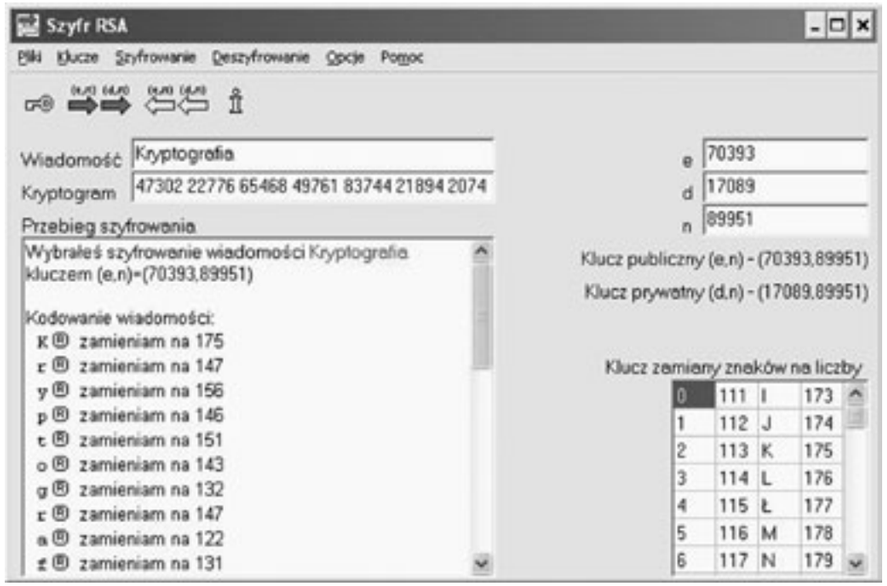
Odszyfrowanie kryptogramu

Bogdan otrzymał od Alicji zaszyfrowaną wiadomość P i aby ją odszyfrować oblicza $M = P^d \pmod{n}$.
Przykład. Mamy obliczyć $M = 16^{103} \pmod{143}$. Postępujemy podobnie, jak w punkcie 2 przy szyfrowaniu wiadomości. Mamy $16^{103} = 16*16^2*16^4*16^{32}*16^{64}$ i z każdej z tych potęg obliczamy resztę z dzielenia przez 143. Otrzymujemy: $16*16^2*16^4*16^{32}*16^{64} = 16*113*42*113*42 = 81 \pmod{143}$.
Bogdan wie, że Alicja chciała jemu przekazać w tajemnicy wiadomość brzmiącą: Q.



Rysunek 6.
Schemat użycia szyfru RSA

Działanie szyfru RSA można prześledzić na przykładach, posługując się w tym celu programem edukacyjnym **Szyfr RSA** (patrz rys. 7). Program ten jest dostępny na stronie programu Informatyka+. Najpierw ustalamy w nim klucz publiczny i klucz prywatny, następnie szyfrujemy wybraną wiadomość, a na końcu odszyfrujemy ją. W głównym oknie programu można śledzić kolejne kroki i obliczenia służące do otrzymywania kluczy, szyfrowania i deszyfrowania wiadomości.



Rysunek 7.
Okno programu edukacyjnego **Szyfr RSA**

Szyfrowanie z kluczem jawnym ma ciekawe zastosowanie. Uczniowie biorący udział w Olimpiadzie Informatycznej w pierwszym etapie zawodów przesyłają rozwiązania na serwer organizatorów konkursu. Organizatorzy chcą zapewnić wszystkich uczniów, że ich rozwiązania dotrą bezpiecznie, to znaczy między innymi, że nie zostaną przechwycone i podpatrzone przez innych uczestników zawodów. Organizatorzy nie wiedzą jednak,

kto będzie startował. W tej sytuacji stosowany jest szyfr RSA – klucz publiczny jest zamieszczony na stronie olimpiady, by mógł z niego skorzystać każdy, kto chce wziąć udział w zawodach i wysłać do organizatorów zadania pierwszego etapu, natomiast klucz prywatny jest przechowywany przez organizatorów, by tylko jury olimpiady mogło zdeszyfrować rozwiązania. A zatem jest możliwe, by ten sam klucz publiczny mógł być stosowany przez wielu nadawców i każdy z nich może być pewien, że po zaszyfrowaniu nim swojej wiadomości będzie mogła ona być odszyfrowana tylko przez osoby, które dysponują prywatnym kluczem z pary – takimi osobami są członkowie jury olimpiady.

Szyfr RSA został wykorzystany w komputerowej realizacji szyfrowania z jawnym kluczem, zwanej **szyfrem PGP** (ang. *Pretty Good Privacy*), który jest powszechnie stosowany w Internecie.

9 PODPIS ELEKTRONICZNY

Jednym z produktów kryptografii jest **podpis elektroniczny** (zwany również **podpisem cyfrowym**), czyli podpis na dokumentach elektronicznych, a raczej – towarzyszący takim dokumentom. O jego znaczeniu może świadczyć fakt, że państwa Unii Europejskiej wprowadzają przepisy, uznające podpis elektroniczny za równorzędny z odręcznym. W Polsce również uchwalono w sejmie odpowiednią ustawę. A więc w przyszłości, zamiast iść do banku, by podpisać umowę o kredyt wystarczy taką umowę zaszyfrować, dołączyć do niej podpis elektroniczny i przesłać pocztą elektroniczną.

Podpis na dokumencie (tradycyjnym lub elektronicznym) jest odpowiedni, jeśli:
- zapewnia jednoznaczną identyfikację autora – nikt inny nie posługuje się takim samym podpisem;
- nie można go podrobić;
- nie można go skopiować na inny dokument – mechanicznie lub elektronicznie;
- gwarantuje, że po podpisaniu nim dokumentu nikt nie może wprowadzić żadnych zmian do tego dokumentu.

Opiszemy teraz, co to jest podpis elektroniczny i jak się nim posługiwać.

Sam podpis elektroniczny jest umownym ciągiem znaków, nierozerwalnie związanych z podpisywanym dokumentem. Jeśli chcemy posługiwać się takim podpisem, to najpierw należy skontaktować się z **centrum certyfikacji**. Tam otrzymuje się parę kluczy, publiczny i prywatny oraz osobisty **certyfikat elektroniczny**, który będzie zawierał m.in. osobiste dane i własny klucz publiczny. Osobisty certyfikat będzie dostępny w Internecie i może być wykorzystany przez odbiorcę dokumentu podpisanego elektronicznie do sprawdzenia tożsamości nadawcy.

Jeszcze jedno pojęcie jest ważne przy posługiwaniu się podpisem elektronicznym, to tzw. **skrót dokumentu**. Skrót powstaje przez zastosowanie do dokumentu odpowiedniego algorytmu (będącego realizacją tzw. **funkcji mieszającej**), np. SHA-1 lub MD-5. Skrót jest jednoznaczną reprezentacją dokumentu, tj. jakakolwiek zmiana w treści dokumentu powoduje zmianę w jego skrócie. Skrót zostaje również zaszyfrowany, jest na stałe związany z dokumentem, na podstawie którego powstał, i identyfikuje podpisującego.

Teraz możemy już opisać, na czym polega podpis elektroniczny dokumentu. Przypuśćmy, że chcę wysłać do Pana A dokument i podpisać go elektronicznie. Dokument znajduje się w komputerze, tam również jest dostępny mój certyfikat elektroniczny. Naciskam więc przycisk **Podpisz**. Wtedy są wykonywane następujące czynności (patrz rys. 8):
1. Tworzony jest skrót z dokumentu przeznaczonego do wysyłki.
2. Skrót ten zostaje zaszyfrowany moim kluczem prywatnym.
3. Do dokumentu zostaje dołączony jego zaszyfrowany skrót i mój certyfikat z kluczem publicznym.
4. Tak podpisany dokument jest przesyłany do odbiorcy, np. pocztą elektroniczną.

Rysunek 8.
Schemat składania podpisu elektronicznego pod dokumentem elektronicznym wysyłanym w sieci
[źródło: http://www.proinfo.com.pl/produkty/podpis-elektroniczny/]

Rysunek 9.
Weryfikacja podpisu elektronicznego pod dokumentem elektronicznym wysyłanym w sieci
[źródło: http://www.proinfo.com.pl/produkty/podpis-elektroniczny/]

Odbiorca, po otrzymaniu dokumentu podpisanego elektronicznie, aby zweryfikować jego autentyczność, naciska przycisk **Weryfikuj**. Wtedy (patrz rys. *9*):

1. Tworzony jest skrót z otrzymanego dokumentu.
2. Dołączony do dokumentu skrót zostaje rozszyfrowany moim kluczem publicznym, zawartym w moim certyfikacie elektronicznym.
3. Utworzony skrót z otrzymanego dokumentu i odszyfrowany skrót zostają porównane. Jeśli są zgodne, to oznacza, że od momentu podpisania dokument nie był modyfikowany oraz że ja, czyli osoba widniejąca na certyfikacie, jestem jego autorem.

Sam dokument może być również zaszyfrowany, np. kluczem publicznym odbiorcy.

Klucz prywatny nadawcy, najbardziej newralgiczna część całego procesu, może być umieszczony na karcie mikroprocesorowej, wtedy, by stosować podpis elektroniczny, należy wyposażyć się w specjalny czytnik. Klucz prywatny, raz umieszczony na karcie, pozostaje na niej cały czas. W takim przypadku, tworzenie skrótu dokumentu z udziałem tego klucza odbywa się również na tej karcie.

Przekonajmy się, że tak tworzony podpis ma wymienione na początku cechy:

- autentyczność nadawcy potwierdza jego certyfikat, do którego ma dostęp odbiorca;
- takiego podpisu nie można podrobić, bo klucz prywatny, pasujący do klucza publicznego znajdującego się w certyfikacie, ma tylko nadawca;
- podpisu nie można związać z innym dokumentem, gdyż nie będzie pasował do jego skrótu;
- skrót dokumentu jest również gwarancją, że dokument nie uległ zmianie po jego podpisaniu, gdyż inaczej skrót odszyfrowany przez odbiorcę nie zgadzałby się ze skrótem, utworzonym przez odbiorcę na podstawie odszyfrowanego listu.

Tworzy się **centra certyfikacji**, które wydają klucze wraz z niezbędnym oprogramowaniem. Co więcej, jeśli odbiorca naszych listów będzie się chciał dodatkowo upewnić, że koresponduje rzeczywiście z daną osobą, to będzie to mógł zrobić, kontaktując się z centrum, które wydało klucze nadawcy. Szczegółowy opis działania i tworzenia podpisu elektronicznego jest opisany na przykład na stronie pod adresem http://www.podpiselektroniczny.pl. Można tam również znaleźć informacje o centrach certyfikacji, jak również utworzyć swój osobisty certyfikat.

Na zakończenie wspomnijmy, że kryptografia ma także swoją ciemną stronę. Szyframi mogą bowiem posługiwać się osoby, które chcą ukryć przed innymi swoje nieczyste zamiary. W ten sposób państwo, które konstytucyjnie ma gwarantować swoim obywatelom bezpieczeństwo, może nie być w stanie wypełnić swojego zobowiązania, gdyby na przykład grupy przestępcze korzystały z szyfrów z kluczami, które zapewniają, że są to szyfry nie do złamania. Dlatego na przykład standardowy szyfr DES, wprowadzony w Stanach Zjednoczonych, ma ograniczoną długość klucza.

Ochrona i bezpieczeństwo przesyłanych wiadomości i przechowywanych danych ma fundamentalne znaczenie dla bezpieczeństwa człowieka i ochrony transakcji, wykonywanych elektronicznie. Z tego powodu kryptografia jest obecnie jednym z najaktywniej rozwijających się działów informatyki i wiele jej osiągnięć zapewne nawet nie jest ujawnianych. Duże nadzieje pokłada się w wykorzystaniu efektów fizyki kwantowej w tworzeniu nowych systemów kryptograficznych. Zainteresowanych szerszymi rozważaniami odsyłamy do wspomnianej już książki Simona Singha, *Księga szyfrów* [8], w której można przeczytać wiele fascynujących historii wiążących się z kryptografią.

10 ALGORYTMY WYKORZYSTYWANE W METODACH SZYFROWANIA

Omawiamy tutaj krótko algorytmy, a faktycznie wymieniamy je tylko, które są wykorzystywane w metodzie szyfrowania RSA. Chcemy pokazać, że algorytmy szyfrowania i deszyfrowania w tej metodzie są bardzo proste i sprowadzają się do wykonania działań, które są przedmiotem zajęć z informatyki w szkole. Niezbędne jest jednak poczynienie pewnych zastrzeżeń:

- uzasadnienie (dowody), że podane algorytmy szyfrowania i deszyfrowania działają poprawnie wykracza poza poziom matematyki szkolnej, nie jesteśmy więc w stanie wykazać, że wiadomość, jaką otrzymuje Bogdan jest dokładnie tą wiadomością, którą wysłała Alicja;
- działania w obu algorytmach, szyfrowania i deszyfrowania, są wykonywane na bardzo dużych liczbach, niezbędne jest więc posłużenie się w obliczeniach komputerowych specjalną arytmetyką długich liczb – nie piszemy o tym tutaj;
- bezpieczeństwo szyfru RSA jest oparte na aktualnym stanie wiedzy informatycznej, zarówno odnośnie problemów, które potrafimy rozwiązywać szybko, jak i problemów, których nie potrafimy rozwiązywać dysponując nawet najszybszymi komputerami.

Skomentujmy więc sposób realizacji poszczególnych kroków algorytmów szyfrowania i deszyfrowania. Komentowane fragmenty tych kroków są poniżej wyróżnione pismem pochyłym.

Przygotowanie kluczy

1. *Bogdan wybiera dwie duże, możliwie bliskie sobie liczby pierwsze p i q. Liczby te powinny mieć przynajmniej po 200 cyfr.*
   Znanych jest wiele metod generowania dużych liczb i sprawdzania, czy są one liczbami pierwszymi. Stosuje się przy tym często metody probabilistyczne, które z prawdopodobieństwem niemal 1 gwarantują, że są to liczby pierwsze.
2. *Bogdan oblicza n = p*q i wybiera dowolną liczbę naturalną e, która jest względnie pierwsza z liczbą (p – 1)*(q – 1), czyli te dwie liczby nie mają wspólnych dzielników poza liczbą 1.*
   Podobnie jak w punkcie 1.
3. *Bogdan znajduje liczbę naturalną d taką, że e*d = 1 (mod (p – 1)*(q – 1)), czyli reszta z dzielenia e*d przez (p – 1)*(q – 1) jest równa 1.*
   Liczbę *d* można znaleźć, posługując się rozszerzonym algorytmem Euklidesa. Algorytm Euklidesa, nawet dla dużych liczb działa bardzo szybko.
4. *Bogdan ogłasza parę liczb (n, e) jako swój klucz publiczny, a parę (n, d) zachowuje w tajemnicy jako klucz prywatny. Jednocześnie niszczy liczby p i q, by nie wpadły w niczyje ręce, co mogłoby umożliwić odtworzenie jej klucza prywatnego.*
   Jeśli *p* i *q* są dostatecznie dużymi liczbami pierwszymi, to znajomość *n* i *e* nie wystarcza, by obliczyć wartość *d* posługując się nawet najpotężniejszymi obecnie komputerami.

Szyfrowanie wiadomości

1. *Jeśli Alicja chce wysłać do Bogdana jakąś wiadomość, to najpierw musi ją przedstawić w postaci liczby naturalnej M, nie większej niż n. Tekst można zamienić na liczbę posługując się kodem ASCII.*
   Jeśli wiadomość jest zbyt długa, to należy ją szyfrować blokami odpowiedniej wielkości.
   To są proste rachunki na reprezentacjach liczb.
2. *Wiadomość, jaką Alicja wysyła do Bogdana, jest liczbą: $P = M^e \pmod n$.*
   Obliczenie wartości *P* można wykonać szybko przy wykorzystaniu szybkiej metody potęgowania, która w pierwszym kroku polega na przedstawieniu wykładnika w postaci binarnej. Wszystkie obliczenia są wykonywane na liczbach nie większych niż *n*. Te obliczenia ilustrujemy w naszym przykładzie przy opisie algorytmu.
3. *Alicja wysyła do Bogdana wiadomość P = 16.*

Odszyfrowanie kryptogramu

Bogdan odszyfrowuje otrzymaną wiadomość P wykonując obliczenia $M = P^d \pmod n$.
Obliczanie wartości *M* jest wykonywane takim samym algorytmem jak *P* w kroku 6.

Na zakończenie naszych rozważań przytoczmy jeden przykład efektywnych obliczeń na dużych liczbach, by uzmysłowić sobie, że obliczeniowa moc komputerów ma swoje ograniczenia i faktycznie cała nadzieja w szybkich algorytmach:

Przypuśćmy, że mamy obliczyć wartość potęgi[21]:

$$x^{12345678912345678912345678912345}$$

której wykładnik ma 32 cyfry. Jeśli zastosujemy szkolną metodę, która polega na kolejnym wykonywaniu mnożeń, to będziemy musieli wykonać ich o jedno mniej niż wynosi wykładnik. Przypuśćmy, że dysponujemy superkomputerem, który wykonuje 10^{15} mnożeń na sekundę. Wtedy ta potęga byłaby obliczona dopiero po $8*10^8$ latach.

A teraz zastosujmy algorytm „binarny", podobny do tego, którego użyliśmy w przykładowych obliczeniach. Wtedy wartość powyższej potęgi zostanie obliczona za pomocą około… 200 mnożeń i będzie to trwało niewielki ułamek sekundy.

LITERATURA

1. Ferguson N., Schneider B., *Kryptografia w praktyce*, Helion, Gliwice 2004
2. Grajek M., *Enigma. Bliżej prawdy*, Rebis, Poznań 2007
3. Gurbiel E., Hard-Olejniczak G., Kołczyk E., Krupicka H., Sysło M.M., *Informatyka, Część 1 i 2, Podręcznik dla LO*, WSiP, Warszawa 2002-2003
4. Harel D., *Algorytmika. Rzecz o istocie informatyki*, WNT, Warszawa 1992
5. Hodges A., *Enigma. Życie i śmierć Alana Turinga*, Prószyński i S-ka, Warszawa 2002
6. Knuth D.E., *Sztuka programowania*, tomy 1–3, WNT, Warszawa 2003
7. Kozaczuk W., *W kręgu Enigmy*, Książka i Wiedza, Warszawa 1986
8. Singh S., *Księga szyfrów. Od starożytnego Egiptu do kryptografii kwantowej*, Albatros, Warszawa 2001
9. Sysło M.M., *Algorytmy*, WSiP, Warszawa 1997
10. Sysło M.M., *Piramidy, szyszki i inne konstrukcje algorytmiczne*, WSiP, Warszawa 1998. Kolejne rozdziały tej książki są zamieszczone na stronie: http://www.wsipnet.pl/kluby/informatyka_ekstra.php?k=69
11. Wobst R., *Kryptologia. Budowa i łamanie zabezpieczeń*, Wydawnictwo RM, Warszawa 2002

[21] Więcej na ten temat – punkt 4.2 w rozdziale Czy wszystko można policzyć na komputerze.