Literatura

[1] *Diffie, W., Hellman, M.*: New directions in cryptography. IEEE Transactions on Information Theory, roč. 22 (1976), č. 6, s. 644–654.

[2] *Knuth, D. E.*: The Art of Computer Programming, Volume 3: Sorting and Searching. Addison-Wesley, 1998.

[3] *Levy, S.*: Crypto: How the Code Rebels Beat the Government Saving Privacy in the Digital Age. Penguin Books, 2001.

[4] *Singh, S.*: Kniha kódů a šifer: Tajná komunikace od starého Egypta po kvantovou kryptografii. Argo, Dokořán, 2003.

Dva základní šifrovací principy

MIROSLAV KOLAŘÍK

Přírodovědecká fakulta UP, Olomouc

Ukážeme si základní principy ze zajímavé a stále se rozvíjející oblasti informatiky – z kryptologie.[1)] Kryptologie zahrnuje kryptografii a kryptoanalýzu. Kryptografie se zabývá metodami utajení (říkáme také šifrováním) obsahu zpráv; toto utajení se provádí tak, že se zpráva převede do podoby, která je srozumitelná pouze zamýšlenému příjemci. Kryptoanalýza se zabývá luštěním zašifrovaných zpráv (dešifrováním).

První kryptografické metody se objevily už ve starověkém Řecku a až do začátku 20. století byly založené na jednoduchých principech. Představíme si dva základní šifrovací principy: substituci a transpozici. Jedná se o jednoduché a přitom zásadní šifrovací metody, které se využívají i u moderních kryptografických metod, jako jsou například DES nebo AES. Šifrovací principy demonstrujeme na několika jednoduchých příkladech, které mohou být použity při výuce na základních a středních školách.

Substituční šifry

Jako první se budeme věnovat substituci, neboli záměně (náhradě). Začneme pěkně zvolna jednoduchým příkladem. Představme si situaci, že

[1)] Kryptologie je naukou o metodách utajování zpráv i o tom, jak zašifrované zprávy luštit (dešifrovat).

chceme zašifrovat text „INFORMATIKA JE SUPER“. Použijeme k tomu následující tabulku, ve které snadno najdeme všechna písmena anglické abecedy, a jak je zašifrovat. Jednoduše, horní písmeno z každého pole tabulky zašifrujeme písmenem ležícím pod ním (ve stejném poli), přičemž mezeru neměníme.

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

Například písmeno I zašifrujeme jako F, podobně písmeno N zašifrujeme jako R a text „INFORMATIKA JE SUPER“ zašifrujeme na text „FRWUYHGXFAG VS TEQSY“. Pro jednoduchost budeme v příkladech vynechávat háčky, čárky a také kroužky. Proto v tabulkách uvádíme pouze písmena anglické abecedy.

Opusťme naši první tabulku a podívejme se na jiný způsob, jak lze nahradit anglická písmena pomocí posloupnosti symbolů, tentokrát tvořených tečkami a čárkami.

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

S využitím druhé tabulky a dohody, že jednotlivá písmena od sebe oddělíme symbolem | (přičemž mezery vynecháme), zakódujeme text „INFORMATIKA“ na tvar „|..|-.|...-.|---|.-.|--|.-|-|..|-.-|.-|“. Pozorný čtenář jistě poznal Morseovu abecedu. Ano, i na morseovku je možné nahlížet jako na substituční šifru.[2)] Toto kódování vymyslel ame-

[2)] Morseovu abecedu však nepoužíváme proto, abychom zprávu utajili, ale pouze *kódovali*, tedy převedli do tvaru, se kterým umí pracovat dané technické zařízení. V případě Morseova kódu je tímto zařízením elektrický telegraf.

rický vynálezce, malíř a sochař Samuel Morse. Ten také uskutečnil v roce 1844 první telegrafické spojení mezi Washingtonem a Baltimorem. Znaky zvolil tak, aby v angličtině nejfrekventovanějším písmenům odpovídaly nejkratší sekvence teček a čárek. Na podobné optimalizaci, kdy nejčastěji používaným znakům je přiřazen nejkratší kód, je založeno také tzv. Huffmanovo kódování, viz např. [1].

Písmena můžeme nahrazovat i nepísmennými symboly, třeba obrázky nebo čísly. Jako ukázka poslouží následující tabulka, ve které jsou písmena anglické abecedy nahrazena jednoduchými obrázky a také je jim přiřazena odpovídající číselná hodnota tzv. ASCII kódu.

A	B	C	D	E	F	G	H	I
♣	☺	★	♯	©	♠	⊗	⇳	∗
65	66	67	68	69	70	71	72	73
J	K	L	M	N	O	P	Q	R
•	⊛	□	☾	≺	♡	≻	Φ	☹
74	75	76	77	78	79	80	81	82
S	T	U	V	W	X	Y	Z	
⊞	⋈	◎	☻	Θ	⇜⇝	§	♢	
83	84	85	86	87	88	89	90	

S využitím této tabulky převedeme slovo „ALGORITMUS“ na „♣□⊗♡☹∗⋈☾◎⊞“ a na číslo „65767179827384778583“. Samozřejmě můžeme nahrazovat i nepísmenné symboly (čísla, mezery, atd.) libovolnými symboly. Však také (volně řečeno) v informatice se pro počítač vše převádí na posloupnosti nul a jedniček.

Jak substituční šifru dešifrovat?

Podívejme se nyní, jak lze jednoduchou substituční šifru dešifrovat. Využijeme toho, že některá písmena se obvykle vyskytují častěji než jiná. Třeba v angličtině se mnohem častěji objeví písmeno E než písmeno Z. My se zaměříme podrobněji na češtinu. Typický výskyt písmen v běžném (dostatečně dlouhém) českém textu rozdělíme do následujících tří skupin:

- nejčastější výskyt (více než pět procent): E, A, O, I, L, N, S, T,
- běžný výskyt (mezi třemi a pěti procenty): D, K, M, P, R, U, V, Z,

- nejméně častý výskyt (méně než tři procenta): Q, X, W, F, G, B, C, H, J, Y.

Dělení je bez diakritiky (nerozlišujeme mezi E, É a Ě, atp.). Nejčastěji se v českém jazyce vyskytují samohlásky E, A a O, všechny s více než osmiprocentní pravděpodobností. Nejméně časté jsou pak souhlásky Q, X, W, F a G s méně než $\frac{1}{4}$procentní pravděpodobností výskytu. Jak lze takové pravděpodobnosti získat? Můžete si to vyzkoušet sami. Sežeňte si český, dostatečně dlouhý text. Čím delší, tím lepší, řekněme, s minimálně desetitisíci znaky. Poté zjistěte výskyty jednotlivých písmen a uvidíte, že to bude hodně podobné s rozdělením výše. Budeme-li nyní dešifrovat nějaký český, jednoduchou substitucí zašifrovaný text, můžeme s výhodou frekvenční analýzu použít. Můžete si to promyslet na konkrétní úloze. Jak bývá zvykem, budeme nyní psát písmena otevřeného (vyluštěného) textu malá a písmena zašifrovaného textu velká.

Úkol 1

Dešifrujte následující text:
ICM ZHVYNAM YJLJYJ YNRYH EN UNZDM UGLJWAN UFMY U IJYSF DNFNCG DNFM ENWAJYZMUGDM VZJUG. ASICMQZSW ENWAJIMVDNAAGKL VZJU EN UNZDM DSZJ, ANFANZN ICNWZJFQG S VIJEQG S S M. VASWAJ YSQ XCFG HLJWANDN ANQYNCS IMVDNAS. WSZN VN DHFNDN FSDNCMY YCNXS AS WUJEMKN IMVDNA S UGHFMY YJLJ, FN ANQYNCN EVJH KSVYN S ANQYNCN VN ANUGVQGYHEM UHXNK. ASVZNWAN KNZQJUN WJZHVYNAM XGUS LJWAN CGKLZN, ICJYJFN ZFN VASWAJ WJIZAJUSY KLGXNEMKM IMVDNAS KNZGKL VZJU.

Řešení. Víme, že jde o český text (bez diakritiky), který byl zašifrován substitucí. Přestože je krátký, získáme frekvenční analýzu všech jednotlivých symbolů. Výsledek absolutní četnosti všech 356 písmenných znaků je znázorněn v následující tabulce.

A	B	C	D	E	F	G	H	I
27	0	12	14	9	13	15	9	11
J	K	L	M	N	O	P	Q	R
27	10	9	21	53	0	0	9	1
S	T	U	V	W	X	Y	Z	
23	0	16	18	14	5	21	19	

Nejčastěji se v šifrovém textu vyskytuje písmeno N, je to tedy horký kandidát na písmeno e v otevřeném textu. Hodně často se vyskytují i písmena A, J a S. Písmeno A se v jednom slově v šifře vyskytuje dvakrát za sebou, nabízí se proto, že by se v otevřeném textu mohlo jednat o písmeno n. Písmeno S se v šifrovém textu vyskytuje několikrát jako samotné jednopísmenné slovo, bude se proto v otevřeném textu nejspíše jednat o písmeno a. Nahradíme-li v šifrovaném textu tato tři písmena a podíváme se na mezivýsledek, snadno odhadneme, že často se vyskytující J bude v otevřeném textu samohláska, nejspíše o. Dostáváme tak slibný mezivýsledek a můžeme pokračovat dále.

Samozřejmě se občas stane, že nějaké písmeno (nějaká písmena) netipneme hned správně. Později ale začnou vycházet nesmysly, což nás přirozeně navede na změnu tak, aby výsledná zpráva dávala smysl. Celý otevřený text vypadá takto: *pri lusteni tohoto textu je velmi vyhodne vzit v potaz mezery mezi jednotlivymi slovy. napriklad jednopismennych slov je velmi malo, neznele predlozky a spojky a a i. snadno tak brzy uhodneme nektera pismena. dale se muzeme zamerit treba na dvojice pismen a vyuzit toho, ze nektere jsou caste a nektere se nevyskytuji vubec. nasledne celkove dolusteni byva hodne rychle, protoze lze snadno doplnovat chybejici pismena celych slov.*

Níže je uvedena tabulka, převádějící písmena šifrového textu (velká) na písmena otevřeného textu (malá). Čtyři písmena nebyla použita (jsou u nich uvedeny otazníky), konkrétně: f, g, q, w, což dobře koresponduje s frekvenční analýzou českého jazyka.

Poznamenejme, že pro ztížení dešifrování substitučních zpráv bývá zvykem psát text bez interpunkce, po pěti znacích oddělených mezerou. Tento zvyk se používá od poloviny 19. století a souvisí s telegrafováním, kde se platilo za slovo, přičemž průměrná délka slov u většiny evropských jazyků je právě pět znaků. Pro názornost, začátek šifrového textu z prvního úkolu by vypadal takto: ICMZH VYNAM YJLJY JYNRY HENUN atd.

A	B	C	D	E	F	G	H	I
n	?	r	m	j	z	y	u	p
J	K	L	M	N	O	P	Q	R
o	c	h	i	e	?	?	k	x
S	T	U	V	W	X	Y	Z	
a	?	v	s	d	b	t	f	

Transpoziční šifry

Dalším základním šifrovacím principem je šifrování pomocí transpozice (přesunu). Opět začneme ukázkovým příkladem.

Mějme text „TRANSPOZICE SPOČÍVÁ VE ZMĚNĚ POŘADÍ ZNAKŮ DLE URČITÉHO PRAVIDLA“. Z daného textu odstraníme interpunkci a mezery a napíšeme jej do tabulky o sedmi sloupcích. Poznamenejme, že počet sloupců můžeme zvolit náhodně (s ohledem na délku textu).

T	R	A	N	S	P	O
Z	I	C	E	S	P	O
C	I	V	A	V	E	Z
M	E	N	E	P	O	R
A	D	I	Z	N	A	K
U	D	L	E	U	R	C
I	T	E	H	O	P	R
A	V	I	D	L	A	

K zašifrování nyní stačí psát zprávu po jednotlivých sloupcích. Dostaneme tak text ve tvaru „TZCMAUIARIIEDDTVACVNILEINEAEZE HDSSVPNUOLPPEOARPAOOZRKCR“, který je na první pohled nesrozumitelný. Provedli jsme transpozici tak, že jsme ponechali původní písmena a jen změnili jejich pozice. Abychom ztížili možnost dešifrovat text, vyzkoušíme psát sloupce v různém pořadí. Použijeme k tomu nějaké sedmipísmenné slovo, řekněme „TELEFON“. Abecední pořadí písmen ve slově „TELEFON“ nám určí, který sloupec budeme psát první, který druhý atd. V našem příkladu budeme psát nejprve druhý sloupec a po něm hned čtvrtý, protože se pojí s písmenem E. Další v abecedě je písmeno F, a proto přepíšeme pátý sloupec, dále je písmeno L a tedy třetí sloupec, atd. Dostaneme tak text „RIIEDDTVNEAEZEHDSSVPNUOLACVNILEIO OZRKCRPPEOARPATZCMAUIA“.

Jinou možností je přepsat písmena z tabulky dle určitého vzoru, třeba podle tvaru spirály. Začneme-li vlevo nahoře a vydáme se směrem dolů a dále pak po obvodu (jak ukazují čísla v následující tabulce), obdržíme text ve tvaru „TZCMAUIAVIDLARCKRZOOPSNARIIEDDTEHOPRAOEP SECVNILEUNPVAEZ“.

1	26	25	24	23	22	21
2	27	44	43	42	41	20
3	28	45	54	53	40	19
4	29	46	55	52	39	18
5	30	47	56	51	38	17
6	31	48	49	50	37	16
7	32	33	34	35	36	15
8	9	10	11	12	13	14

Napsání textu pozpátku je dalším jednoduchým příkladem transpoziční šifry. Nebo třeba psaní nejprve písmen na lichých pozicích, pak na sudých pozicích. Příkladem je také tzv. skytalé. Tento historický šifrovací princip využívá válce a pásku papíru (dříve pergamenu). Páska papíru se postupně namotá na válec a až je namotaná napíše se na ni (ve směru osy válce) vzkaz. Po odmotání jsou písmena na pásce přemístěna a text se stává (na první pohled) nesrozumitelným. Nejrychleji se zpráva vyluští namotáním na válec o stejném průměru, jako měl původní válec. K dešifrování lze použít i kužel, na který se páska namotá, a poté se po něm posouvá do té doby, než se objeví smysluplné slovo. Tento druh šifrování používali Sparťané během válek.

Další hodně jednoduchou šifru, která přepisuje písmenka „cik-cak“ si demonstrujeme na tabulce s dvěma řádky a jedenácti sloupci. Samozřejmě pro jiné rozměry není těžké šifrovací algoritmus adekvátně pozměnit.

1	21	3	19	5	17	7	15	9	13	11
22	2	20	4	18	6	16	8	14	10	12

Zprávu nejprve přepíšeme po řádcích. Čísla v tabulce udávají pořadí přepisu jednotlivých písmen zprávy. Budeme-li chtít zašifrovat konkrétní zprávu „ENIGMA JE SIFROVACI STROJ“, tak ji nejprve po řádcích zapíšeme do tabulky.

E	N	I	G	M	A	J	E	S	I	F
R	O	V	A	C	I	S	T	R	O	J

Poté text dle „cik-cak“ postupu jednoduše přepíšeme na „EOIAMIJT SOFJIRESACGVNR“ a zašifrovaný text je hotový.

Jako poslední jednoduchou ukázku transpoziční šifry uvedeme zašifrování textu „ENIGMA JE STROJ“ pomocí tabulky o rozměrech 13×4 polí. Text napíšeme předem dohodnutým způsobem, například jako v následující tabulce, a poté jej přepíšeme po jednotlivých řádcích.

E						J						J
	N				A		E				O	
		I		M				S		R		
			G						T			

V našem případě obdržíme šifrový text ve tvaru „EJJNAEOIMSRGT“.

Otočná mřížka

Nyní se zaměříme na tzv. Fleissnerovu otočnou mřížku. Tuto transpoziční šifru popsal jako první Fleissner von Wostrowitz v roce 1881. K šifrování a dešifrování se používá pomůcka: čtvercová tabulka, ve které jsou na vybraných místech vystřihnuty malé čtverce. Například v následující tabulce velikosti 5×5 jsme čísly vyznačili vystřižené (malé) čtverce.

1				
				2
			3	
	4			5
		6		

Čtverce se vystřihují tak, aby se při postupné rotaci celého velkého čtverce (doprava o 90, 180 a 270 stupňů) vystřižené (malé) čtverce nikdy nepřekryly. Text zprávy se pak postupně vpisuje do vystřižených (malých) čtverců a po jejich zaplnění se mřížka otočí o 90 stupňů směrem doprava. S naší ukázkovou otočnou mřížkou tak můžeme zašifrovat zprávu o 24 symbolech. Jejich umístění znázorňují čísla 1 až 24 v následující tabulce. Prostřední políčko zůstalo prázdné a nebude se využívat. Šifrový text vytvoříme jednoduše tak, že jej čteme po jednotlivých řádcích.

1	19	13	20	7
14	8	21	15	2
9	16		3	22
17	4	10	23	5
24	11	6	12	18

Ukažme si postup na konkrétním příkladě. Chceme zašifrovat zprávu ve tvaru „NOTEBOOK – PŘENOSNÝ POČÍTAČ“. Použijeme k tomu výše uvedenou pomocnou otočnou mřížku velikosti 5×5. Po doplnění textu (bez mezer a diakritiky) obdržíme tuto tabulku:

N	O	N	C	O
O	K	I	S	O
–	N		T	T
Y	E	P	A	B
C	R	O	E	P

Nyní již lehce vytvoříme šifrový text. Přepsáním po jednotlivých řádcích dostáváme „NONCOOKISO–NTTYEPABCROEP“.

Pochopení principu otočné mřížky si můžete ověřit na následujícím úkolu. Šifrový text vznikl pomocí Fleissnerovy otočné mřížky velikosti 4×4. Napovíme, že v každém řádku mřížky je právě jeden vystřižený (malý) čtverec.

Úkol 2

Získejte heslo z následujících 16 písmen: „EOHSBLIEOMSJLVET“.

Řešení. Víme, že šifrový text vznikl pomocí Fleissnerovy otočné mřížky velikosti 4×4. Z nápovědy také víme, že v každém řádku mřížky je právě jedno vystřižené políčko. Uspořádáme-li šifrový text do mřížky obdržíme následující:

E	O	H	S
B	L	I	E
O	M	S	J
L	V	E	T

V prvním řádku bude počáteční písmeno otevřeného textu. Víme, že ve druhém řádku najdeme druhé písmeno otevřeného textu, ve třetím řádku bude třetí písmeno otevřeného textu a ve čtvrtém řádku najdeme čtvrté písmeno otevřeného textu. Otevřený text má být smysluplný, což nám hodně slepých cest eliminuje. Například vybereme-li E z prvního řádku, nebudeme vybírat E z druhého řádku, protože by otevřený text začínající písmeny „EE“ nedával smysl. Podobně můžeme s velkou jistotou okamžitě vyloučit trojice „OIJ“, „SLJ“, „EBM“. Při určování prvních čtyř písmen otevřeného textu je určitě užitečné pracovat s principem šifry – s otáčením mřížky vždy po sepsání čtveřice písmen. Například, vybereme-li jako první písmeno otevřeného textu písmeno O, nebo H, tak víme, která další tři písmena nebudou na druhé, třetí a čtvrté pozici v otevřeném textu. Podaří-li se nám správně určit první čtyři písmena, máme celý klíč k šifře, neboť tak známe polohu všech čtyř vystřižených políček v naší mřížce. Po těchto úvahách lze poměrně rychle získat otevřený text: „HESLEMJESLOVOBIT“. Správnou odpovědí je tedy slovo „BIT“.

Závěr

Představili jsme si dva základní šifrovací principy: SUBSTITUCI při které se znaky otevřeného textu nahrazují šifrovými symboly a TRANSPOZICI, kde se mění pořadí znaků otevřeného textu. Oba tyto principy se používají i v dnešní době – jejich vhodnou kombinací vznikají silné šifry. Dalším šifrovacím principem je například steganografie (ukrývání zpráv). Úkolem steganografie je skrýt samotnou existenci zpráv. Ke steganografickým metodám patří použití neviditelného inkoustu, nenápadné vyznačení písmen v jinak nezávadném textu, ukrytí znaků otevřeného textu na domluvených pozicích, ukrytí zprávy v obrázku na internetu, použití mikroteček apod. Pomocí počítače lze informaci ukrýt, například do obrázku, do audiosignálu, do videa. Ale o tom třeba někdy příště.

Literatura

[1] *Bartl, E.*: Teorie informace, MFI, roč. 24 (2015), č. 3, s. 219–228.