Dětská univerzita a šifrování

Jaroslav Hora

Na pedagogické fakultě ZČU se již delší dobu zúčastňuji výuky v tzv. Dětské univerzitě. Jde o soubor kurzů z mnoha různých oborů, které mají zaujmout talentované žáky zhruba ve věku 10 – 14 let. Výuka probíhá v odpoledních hodinách a v kursu O prvočíslech a šifrování obvykle vyřešíme nějakou náročnější slovní úlohu vyžadující nalezení rozkladu jistého přirozeného čísla v součin prvočísel, realizujeme hledání prvočísel metodou Eratosthenova síta a povíme si něco o historii šifrování, vyřešíme jednu šifru pocházející od E. A. Poea a řekneme si také, že moderní šifrování je založeno na využití velkých prvočísel.

JAK STARÝ BYL KAPITÁN?

„Za první světové války se strhla nedaleko malého městečka bitva. Bylo to právě posledního dne v měsíci. V této bitvě byla rozbita socha bojovníka s píkou ze starých dob. Socha byla asi v životní velikosti.

Násobíme-li datum dne, kdy se bitva strhla, délkou píky ve stopách, polovičním stářím sochy a věkem velícího kapitána, dostaneme číslo 451 066.

K tomu poznamenáváme, že se tehdy dělaly píky dlouhé asi 2 metry.“

Řešení:

Rozložme číslo 451 066 v součin prvočísel:

$$451\,066 = 2 \times 225\,533 = 2 \times 7 \times 32\,219 = 2 \times 7 \times 11 \times 2929 = 2 \times 7 \times 11 \times 29 \times 101.$$

Jako datum posledního dne v měsíci připadá v úvahu jen 29. Oním měsícem musí být únor přestupného roku a tím je během první světové války rok 1916. Socha tedy byla rozbita 29. 2. 1916.

Dále je jasné, že délka píky byla 7 stop, tj. cca 210 cm.

Poloviční stáří sochy je 101 let, tj. socha je stará 202 let.

Věk velícího kapitána je 22 let.

1. ERATOSTHENOVO SÍTO

Začněme jednou klasickou úlohou, známou již ze základní školy.

Příklad: Nalezněte všechna prvočísla p, pro něž platí $1 < p < 120$.

Řešení: Využijeme výše zmíněné Eratosthenovo síto. Napišme (lineární zápis) posloupnost všech přirozených čísel od 1 do 120:

1 2 3 4 5 6 … 118 119 120.

Ideou je vyškrtat „nevhodná“ složená čísla. Číslo 1 není prvočíslem, je tzv. jednotkou ve smyslu dělitelnosti. Následující číslo 2 prvočíslem je. Jeho všechny násobky větší než 2 v této posloupnosti jsou již nezbytně složenými čísly; vyškrtneme je. Následujícím nevyškrtnutým číslem po číslu 2 je číslo 3, které je prvočíslem. Opět vyškrtáme všechny (dosud nevyškrtnuté) násobky čísla 3 větší než toto číslo. Postoupíme na další dosud nevyškrtnuté číslo, tj. na číslo 5. Nalezli jsme další prvočíslo a vyškrtávání opakujeme. Poté přejdeme na prvočíslo 7 a celý postup opakujeme. Smysl slova „síto“ je teď zřejmý: doslova „prosíváme“ přirozená čísla od 1 do 120 (horní mez byla zvolena víceméně náhodou) a získáváme prvočísla.

Kdy ukončit vyškrtávání při realizaci Eratosthenova síta? Snadno se nahlédne, že vyškrtávání v posloupnosti přirozených čísel od 1 do jistého k již není třeba provádět pro žádné $a \in \mathrm{N}$, které je větší než odmocnina z k. V našem případě to znamená, že v posloupnosti obsahující čísla od 1 až do $k = 120$ již není třeba provádět vyškrtávání pro $a = 11$. (V daném případě je číslo $11^2 = 121$ již mimo oblast našeho zájmu a všechny předchozí násobky tvaru jedenácti již byly vyškrtnuty). To znamená, že naše práce skončila vyškrtáváním pro hodnotu $a = 7$! Nechceme zde zabírat zápisem Eratosthenova síta příliš místa, nalezená prvočísla jsou tučně zvýrazněna v tab. 1. níže.

Zapisovat posloupnost všech přirozených čísel od 1 do 120 je nezajímavá, nudná, rutinní práce. Tím bych inteligentní žákovské frekventanty Dětské univerzity příliš neuchvátil. Proto jsem pro ně předem vytvořil tabulku o šesti sloupcích. Uspořádání původně

lineárního seznamu čísel do tabulky je zdánlivě drobná změna. Uvidíme, k čemu to povede. Protože jsme houbařský národ, další postup motivuji následovně.

1	**2**	**3**	4	**5**	6
7	8	9	10	**11**	12
13	14	15	16	**17**	18
19	20	21	22	**23**	24
25	26	27	28	**29**	30
31	32	33	34	35	36
37	38	39	40	**41**	42
43	44	45	46	**47**	48
49	50	51	52	**53**	54
55	56	57	58	**59**	60
61	62	63	64	65	66
67	68	69	70	**71**	72
73	74	75	76	77	78
79	80	81	82	**83**	84
85	86	87	88	**89**	90
91	92	93	94	95	96
97	98	99	100	**101**	102
103	104	105	106	**107**	108
109	110	111	112	**113**	114
115	116	117	118	119	120

Tab.1

Když v průběhu houbařské sezóny vyrazíme do lesa na houby, může se stát, že potkáme dobrého kamaráda, který se již vrací s úlovkem domů. Na otázku, kde rostou, dostaneme často odpověď typu: „V Kovářikovic lesíku nic není, je tam moc sucho. Ani v Soudným nejsou, ale dost rostou v mlází v Nejdlovic lesíku“. S hledáním prvočísel v tab.1 je to podobné. Podívejme se na čísla vyskytující se ve druhém sloupci – mají tvar $6n + 2$, $n = 0 , 1, \dots$. Jde tedy vesměs o sudá čísla a tento sloupec je na prvočísla „neúrodný“, vyskytuje se tu jediné, a to číslo 2. Další sloupec obsahuje čísla tvaru $6n + 3$, $n = 0, 1, \dots$. Máme tu násobky tří a je v něm jediné prvočíslo, totiž 3.

Ve sloupcích obsahujících prvky tvaru $6n + 4$, $n = 0, 1, \dots$ a $6n$, $n = 1, \dots$ zjevně nenalezneme ani jediné prvočíslo. Kdybychom tedy realizovali Eratosthenovo síto právě v této tabulce, mohli bychom si významně ušetřit práci: tyto sloupce bychom zcela vyškrtli a v dalších dvou je pouze jediné prvočíslo na první pozici v daném sloupci. Trocha tvořivého přístupu přinesla první plody, úsporu práce s vyškrtáváním. Povšimněme si, že šlo o sloupce, v nichž byla čísla tvaru $6n + m$, kde největší společný dělitel $D(6, m)$ čísel 6 a m je větší než 1, neboli čísla 6 a m jsou soudělná.

Dokončením nám již známého postupu zjistíme, že se prvočísla soustředila do dvou sloupců tab. 1, totiž do těch, v nichž jsou prvky tvaru $6n + 1$, $n = 0, 1, \dots$ a $6n + 5$, $n = 0 , 1, \dots$. Naše hledání zasáhlo jen přirozená čísla od 1 do 120. Připomeňme, že prvočísel je nekonečně mnoho. Kdybychom nyní uvažovali množiny všech přirozených čísel $6n + 1$, $n = 0 , 1, \dots$ a $6n + 5$, $n = 0, 1, \dots$ bez omezení onou horní hranicí, tj. číslem 120, pak by bylo jasné, že v jejich sjednocení leží nekonečně mnoho prvočísel. Platí však, že v jedné každé z těchto množin je nekonečně mnoho prvočísel?

Popřemýšlejme nyní nad další souvislostí. Co kdyby naše tabulka měla kupř. deset sloupců? Daly by se nějaké vyškrtnout? Nu ano, sloupce, v nichž jsou čísla tvaru $10n + 2$, $10n + 4$, $10n + 5$, $10n + 6$, $10n + 8$, $n = 0, 1, \dots$a $10n = 10n + 0$, $n = 1, \dots$ jsou těmi na prvočísla neúrodnými „lesíky“, v nichž lze nalézt nejvýše jedno prvočíslo. Kdybychom si připravili příslušnou tabulku, viděli bychom, že se „skoro všechna“ prvočísla nacházejí ve sloupcích, jejichž prvky mají tvar $10n + 1$, $10n + 3$, $10n + 7$, $10n + 9$, $n = 0, 1, \dots 11$. Odbouráme vcelku náhodné omezení na čísla menší než 120 a otázka je nasnadě: Je v každé ze tříd $10n + 1$, $10n + 3$, $10n + 7$, $10n + 9$, $n \in \mathrm{N}$ nekonečně mnoho prvočísel? Někteří šikovní frekventanti Dětské univerzity již vědí, že prvočísel je nekonečně mnoho. Na otázky o zaplnění oněch nekonečných „lesíků“ by možná odpověděli správně, vedeni spíše touhou po

„spravedlivém“ rozložení prvočísel. Dostali jsme se ale k obtížnému problému z teorie čísel. Co o něm říci ve výuce algebry jejich budoucím učitelům?

2. HLUBOKÝ PROBLÉM V TEORII ČÍSEL

Poněkud netradiční uspořádání Eratosthenova schématu nás přivedlo k otázce, kterou je možno zobecnit. Nechť a a b jsou dvě nesoudělná přirozená čísla. Existuje v každé aritmetické posloupnosti tvořené prvky tvaru $a\,n + b$, $n \in \mathrm{N}$ nekonečně mnoho prvočísel?

Je dobře známo, že v teorii čísel lze mnohdy zformulovat jednoduše motivované otázky, jejichž řešení je velmi obtížné. To je i právě uvedený případ. Ano, v každé výše popsané aritmetické posloupnosti vskutku existuje nekonečně mnoho prvočísel. Důkaz tohoto tvrzení je náročný a byl nalezen německým matematikem Johannem Peterem Dirichletem (1805 – 1859) v roce 1837. Dirichlet je mj. i proto pokládán za zakladatele tzv. analytické teorie čísel. Jenže jeho důkaz do běžné přípravy učitelů nebudeme moci zařadit, neboť využívá teorii funkcí komplexní proměnné, s níž dnes již nejsou studenti učitelství matematiky pro ZŠ seznamováni.

V roce 1949 podal A. Selberg elementární důkaz Dirichletovy věty o prvočíslech v aritmetických posloupnostech. Slovu elementární je třeba rozumět tak, že se v důkazu nevyužívají funkce komplexní proměnné, ale nejde o důkaz jednoduchý a pro studenty učitelství s matematikou bych jej do běžné výuky nezařadil. Jako vhodný postup bych uvažoval alespoň o předvedení důkazu pro některé speciální hodnoty a, b nesoudělných přirozených čísel.

Cvičení: Dokažte, že v aritmetické posloupnosti $4n + 3$, $n \in \mathrm{N}$ existuje nekonečně mnoho prvočísel.

Nástin důkazu: Předpokládejme, že prvočísel dávajících při dělení 4 zbytek 3 je jen konečně mnoho. Nechť jsou to čísla 3, p_1, p_2, …, p_r. Utvořme číslo $A = 4\,p_1\,p_2 \ldots\, p_r + 3$ a zapišme je jako součin prvočísel ve tvaru $A = q_1\,q_2\,\ldots\,q_s$. Nahlédněme, že aspoň jedno z prvočísel q_1, q_2, ..., q_r dává při dělení 4 zbytek 3. (Kdyby ne, pak jsou všechna tvaru $4x + 1$ a totéž by platilo o jejich součinu A).

Buď tedy q_k takové prvočíslo, které dává při dělení čtyřmi zbytek 3. Víme, že q_k dělí A, ale z původní definice A vidíme, že žádné z čísel 3, p_1, $p_{2,\ldots,}\, p_r$ nedělí A. Tudíž q_k není rovno

žádnému členu v tomto seznamu, jde o nové prvočíslo a dostáváme spor s tím, že prvočísel dávajících při dělení 4 zbytek 3 je jen konečně mnoho.

Ještě poznamenejme, že přemýšlivý a tvořivý žák či student se může od školního Eratosthenova síta dostat až k základům programování, resp. k vytvoření drobného matematického programu. Stačilo by využít kupř. Excelu a namísto vyškrtávání složených čísel je přepisovat např. číslem 0. Nenulová čísla nacházející se v tabulce budou hledanými prvočísly.

3. O ŠIFRÁCH A ŠIFROVÁNÍ

Problém, jak zašifrovat zprávu tak, aby jí nepřítel neporozuměl a spojenec ji opět mohl rozšifrovat, je ve vojenství důležitý odpradávna. Uvádí se, že Gaius Julius Caesar užíval jednoduchou šifru, v níž při šifrování bylo každé písmeno nahrazeno jiným, nacházejícím se v latinské abecedě „o tři místa dále“, tj. A $\rightarrow$ D, B $\rightarrow$ E, C $\rightarrow$ F, ..., Y $\rightarrow$ B, Z $\rightarrow$ C. Známý Caesarův výrok „ ALEA IACTA EST “, „kostky jsou vrženy“, bychom v abecedě o znacích A B C D E F G H I J K L M N O P Q R S T U V X Y Z zašifrovali jako „DOHD LDFXD HVX“. (Původní latinská abeceda měla jen 23 písmen, dnešní znaky J a U chyběly, znaku I se užívalo pro i i pro j, V pro u i pro v, ale to pro pochopení principu šifrování není podstatné). Je jasné, jak získanou zprávu dešifrovat. Zároveň je ale zřejmé, že obdobný šifrovací systém lze snadno prolomit. Výše zmíněné „posunutí o tři písmena“ je opravdu dosti naivní.

Dnešní šifrovací metody budou zajisté muset být podstatně rafinovanější, bude třeba počítat s útoky vysoce kvalifikovaných specialistů majících k dispozici výpočetní techniku. Přibližme se dnešku i tím, že přejdeme k anglické abecedě o 26 písmenech doplněním znaku W. Společně s přítelem se dohodneme, že pro účel předání tajné zprávy o nejvýše n písmenech vygenerujeme náhodnou posloupnost n písmen z naší abecedy. Bude -li kupř. n = 15, vytvoříme třeba kód F K C T Y G M P S B W Z H A Q. Při kódování zprávy „ ALEA IACTA EST “ nyní postupujme následovně. První písmenko naší zprávy je shodou okolností i první v abecedě, první písmeno tajného kódu, tj. F, je v anglické abecedě šesté. Součet 1 + 6 = 7, proto zašifrovaná zpráva bude začínat sedmým písmenem abecedy, tedy G. Druhému písmenu zprávy, tedy L, odpovídá 12, druhému písmenu klíče, tj. K, odpovídá 11, součet činí 23 a tomu odpovídá v anglické abecedě písmeno W. Dále písmena E, resp. C mají čísla 5, resp. 3, obdržíme tedy součet 8 a tomu odpovídá písmeno H.

Idea šifrování je v tomto případě jasná, jen poznamenejme, že v případě, kdy součet vyjde větší než 26, je nutné jej o 26 zmenšit. Obdržíme posloupnost „GWHU HHPJT GPT“.

Příjemce, který zná tajný klíč, zprávu „písmeno po písmenu“ dešifruje (namísto sčítání použije ovšem operaci odčítání). Kupř. tedy prvnímu písmenu zašifrované zprávy, tj. písmenu G, odpovídá číslo 8, prvnímu písmenu klíče, tj. písmenu F, odpovídá 7 a rozdíl 8 – 7 = 1 bude dešifrován jako písmeno A. Dokončete sami dešifrování zprávy!

Výhodou popsaného klíče je jeho bezpečnost, pokud se ovšem podaří utajit šifrovací klíč. Je -li ale tento klíč rozsáhlý, nemusí být utajení snadné. Stane -li se, že zasílaná zpráva

obsahuje více písmen než je dohodnutá délka zprávy *n*, lze sice kód použít znovu, ale toto opakování již poskytuje doplňkové informace i nepříteli a uvádí se, že bezpečnost kódu může být ohrožena. Lze tedy říci, že popsaný „tajný kód na jedno použití“ je bezpečnou, ale poněkud nepohodlnou metodou šifrování.

Jako kulturně – historickou zajímavost uveďme, že luštění šifer tvoří podstatnou část některých povídek zakladatelů moderní detektivky. Jde kupř. o příběhy Tančící figurky z knihy Návrat Sherlocka Holmese či Gloria Scottová z knihy Vzpomínky na Sherlocka Holmese A. C. Doyla. Snad ještě zajímavější je povídka Zlatý skarabeus „otce moderní detektivky“ Edgara Allana Poea (1809 – 1849). Povídka pochází z roku 1843 a v originále se nazývá The Gold–Bug. Klíčem k nalezení pokladu kapitána pirátů Kidda je rozluštění následující šifry:

53++!305))6*;4826)4+.)4+);806*;48!8`60))85;**;**]8***;**:+*8!83(88)5*!;
46(;88*96*?;8)*+(;485);5*!2:*+(;4956*2(5*-4)8`8*;4069285);)6
!8)4++;1(+9;48081;8:8+1;48!85;4)485!528806*81(+9;48;(88;4(+?3
4;48)4+;161;:188;+?;

„V angličtině je nejčastějším písmenem e. Další pořadí je toto: a o i d h n r s t u y c f g l m w b k p q x z. e má takovou převahu, že se sotva najde věta jakkoliv dlouhá, ve které by toto písmeno nepřevládalo.“

To přivede hrdinu povídky k předpokladu, že 8 je třeba nahradit písmenem e. (Budeme -li si chtít luštění šifry v pohodlí zrekapitulovat, můžeme dnes okopírovat zašifrovaný text a uložit jej ve Wordu a pak jednotlivé znaky nahrazovat wordovským povelem Zaměnit z menu Úpravy). Nahradíme -li tedy 8 písmenkem e, obdržíme

53++!305))6*;4e26)4+.)4+);e06*;4e!e`60))e5;;]e*;:+*e!e3(ee)5*!;
46(;ee*96*?;e)*+(;4e5);5*!2:*+(;4956*2(5*-4)e`e*;40692e5);)6
!e)4++;1(+9;4e0e1;e:e+1;4e!e5;4)4e5!52ee06*e1(+9;4e;(ee;4(+?3 4;4e)4+;161;:1ee;+?;

Další luštění šifry provádím v rámci výuky v Dětské univerzitě a uvidíme je v rámci přednášky.

Lze říci, že až do začátku sedmdesátých let byla kryptografie záležitostí povýtce vojenskou, související i se špionáží a diplomatickými službami. O teorii kódů se mnoho nepublikovalo, pokud máme na mysli obvyklé vědecké časopisy, sborníky konferencí atd. Ovšem civilní obchodní a bankovní sféra začala stále více využívat bezdrátový přenos dat. Mohlo jít i o citlivá data o společnostech i jednotlivcích. Nahlédlo se, že by byl velice potřebný jednoduchý kryptografický algoritmus, použitelný pro bezpečný a rychlý přenos dat v mnoha různých podmínkách. V listopadu 1976 byl v USA formálně přijat jako federální standard tzv. DES (Digital Encryption Standard). Šlo o šifrovací systém s tajným klíčem. Libovolní dva uživatelé si musejí tento klíč vyměnit ještě předtím, než si začnou vyměňovat zašifrované zprávy. Samozřejmě, ideou je, že zašifrovanou zprávu bude velice těžké rozluštit pro každého, kdo klíč nezná. Vlastník klíče však zprávu rozšifruje snadno.

Někdy však není jednoduché „rozdat“ tajný klíč všem lidem, kteří ho budou využívat. V roce 1976 navrhli Diffie a Hellman principiálně novou metodu veřejného klíče pro šifrovací

systémy. Problém distribuce klíče se obešel velice elegantně – může jej znát kdokoli. Veřejný klíč je možné kupř. zveřejnit v novinách. Systém však má i tajný klíč. Jen ten, kdo je zná tento tajný klíč, je schopen lehce rozšifrovat text, zašifrovaný veřejným klíčem.

Skvělá idea, ale jak ji implementovat? To navrhli brzy poté Rivest, Shamir a Adleman (podle prvních písmen jejich jmen se hovoří o RSA systému). Jejich metoda je vysoce aktuální i v souvislosti s problematikou elektronického podpisu, obchodování na Internetu atd. Přitom se opírá o záležitosti ryze matematické, kupř. rozkládání (faktorizace) přirozených čísel na mocniny prvočísel a jiné elementární poznatky teorie čísel.

Jak metoda RSA funguje? Nejprve je zapotřebí převést zprávu z běžného jazyka do posloupnosti čísel, což lze, jak již víme, snadno učinit. Obdržíme jakési číslo x. Nyní je třeba vzít dvě „veliká“ navzájem různá prvočísla p, q. Jejich součin $p.q$ je pochopitelně rozumné spočítat na počítači. Tento součin nebude nijak utajován, lze jej kupř. zveřejnit v novinách.

Ten, kdo zná obě prvočísla p a q, může velice snadno spočítat hodnotu Eulerovy funkce $\varphi(p.q) = (p - 1)\,(q - 1)$. Nyní zvolí nějaké číslo e nesoudělné s $\varphi(p.q)$, nikoli však $e = 1$ nebo $e = (p - 1)\,(q - 1) - 1$. Toto číslo (šifrovací exponent) se též zveřejní. Každý, kdo zná čísla $p.q$ a e, může své sdělení zašifrovat (proto se mluví o metodě veřejného klíče). Provede to tak, že vypočte (ovšemže opět s pomocí počítače) číslo y, pro něž platí $y \equiv x^e \pmod{p.q}$. Autor kódu si musí vypočítat ještě jedno číslo (dekódovací exponent f, „tajný“ klíč). Nalezne jej jako řešení kongruence $e.f \equiv 1 \pmod{\varphi(p.q)}$.

To není obtížný úkol, neboť čísla e, $\varphi(p.q)$ jsou, jak řečeno, nesoudělná, D(e, $\varphi(p.q)$) = 1. Hledání čísla f se opírá o tzv. rozšířený Euklidův algoritmus. Z něj plyne, že existují celá čísla f, h taková, že $e.f + \varphi(p.q).h = 1$. Je pak $e.f = 1 - \varphi(p.q).\,h$, čili $e.f - 1$ je dělitelné $\varphi(p.q)$ a $e.f \equiv 1 \pmod{\varphi(p.q)}$.

Nyní již můžeme dešifrovat. Předpokládejme, že pro původní číslo x platí, že $x < p.q$ (kdyby to nebylo splněno, bylo by možné původní zprávu rozdělit na více částí) a že ani p, ani q nedělí x (prakticky vzato, tento příklad pro „veliká“ prvočísla p, q nastane „málokdy“ a problém lze vždy obejít malou úpravou x, která nemění smysl zprávy).

Jak víme, platí $e.f = 1 - \varphi(p.q).\,h$, nebo, pokud označíme $k = -h$, $e.f = 1 + \varphi(p.q).\,k$. Víme dále, že čísla x a $p.q$ jsou nesoudělná, takže podle Eulerovy věty

$$x^{\varphi(p.q)} \equiv 1 \pmod{p.q}$$

a tedy též

$$x^{k.\varphi(p.q)} = \left(x^{\varphi(p.q)}\right)^k \equiv 1 \pmod{p.q}$$

a nakonec $y^f = x^{e.f} = x.\,x^{k.\varphi(p.q)} \equiv x \pmod{p.q}$.

Poslední vztah popisuje, jak zakódované slovo y dešifrovat – postačí vypočítat nejmenší nezáporný zbytek při dělení mocniny y^f modulem $p.q$. To je, jak víme, efektivně řešitelná úloha.