Rozjímání o Rychlé šifře

Podezíráme matku Přírodu, že do DNA zakódovala tři základní šifry: substituci, transpozici a aditivní šifru. Vývoj v uplynulých 4000 letech, kterými prošla věda o utajování, zapříčinil, že tato skutečnost vyplynula na povrch. Lidé od počítačů (říkejme stručně programátoři) tíhnou k té poslední, a konkrétně k její variantě, tzv. Vigenerově šifře, neboť když přitlačíte, aby nějakou nezávaznou šifru rychle vymysleli, přijdou právě s touto šifrou.

Tabulka 1 Caesarova šifra

ABCDEFGHIJK....XYZ
DEFGHIJKLMN...ABC

Vigenerova šifra

U programátorů občas vzniká situace, že píší nějakou aplikaci, kde zpočátku nejde o utajení, ale jak se dílo chýlí ke konci, ukazuje se, že data leží na disku jen tak, a i když se nejedná o primární hrozbu, bylo by zbytečné, aby někoho dráždila. Proto vznikne dodatečný požadavek „nějak to zašifrujte, ale ať to v žádném případě nebrzdí aplikaci“. Tvůrce se nyní soustředí na šifru. Zjistí, že standard nechce použít, protože by to dost zdržovalo. A pak, on potřebuje z daného místa přečíst třeba čtyři bajty a někam jinam uložit třeba padesát bajtů, kdežto AES pracuje s bloky o 16 bajtech.

(Teď maličko přerušíme tok jeho myšlenek. Tohle všechno věda o šifrování už vyřešila tzv. mody šifrování – třeba čítačový modus, modus CFB, proudovými šiframi apod., ale až v posledních třiceti letech, což matka Příroda ještě nestačila zakódovat do našich hlav, zatímco Vigenerovu šifru do nás pěchuje už 500 let.)

Navíc to musí být něco, co velmi rychle naprogramuje. A právě teď (nevíme proč) mozek dodává informaci, že se na data může načítat periodicky opakovaný tajný klíč. To je rozhodující okamžik, dál už je jen milion jeho variant a my budeme rozvíjet pouze jednu. Celé zašifrování navrhne tak, že bude mít v zásadě tvar z = d xor k, kde d jsou tekoucí vstupní data a k je tekoucí klíč. Klíč má délku D a řekněme, že je uložen jako pole bajtů k[0, ..., D-1], které se používá periodicky za sebou. Přesněji má potom algoritmus tvar znak[pozice] = data[pozice] xor k[pozice modulo D]. Výhoda je, že bajty na 1.,2., ..., D- té pozici se pokaždé zakrývají jiným bajtem klíče. Dále tu máme vynikající vlastnost, že algoritmus pro zašifrování je stejný jako algoritmus pro dešifrování! (a tak lze zařadit jen jakoby „driver“ mezi čtení/zápis dat). Skutečně, ať dešifrujeme nebo šifrujeme, vždy provádíme stejnou operaci – xorujeme na vstupní data odpovídající bajt klíče. Navíc, tato operace opravdu nic nezdrží, takže úloha je úspěšně vyřešena. Zbývá ještě klíč, ale o tom až později.

Tabulka 2 Tabula recta

ABCDEFGHIJKLMNOPQRSTUVWXYZ
ABCDEFGHIJKLMNOPQRSTUVWXYZ
BCDEFGHIJKLMNOPQRSTUVWXYZA
CDEFGHIJKLMNOPQRSTUVWXYZAB
DEFGHIJKLMNOPQRSTUVWXYZABC
.....
WXYZABCDEFGHIJKLMNOPQRSTUV
XYZABCDEFGHIJKLMNOPQRSTUVW
YZABCDEFGHIJKLMNOPQRSTUVWX
ZABCDEFGHIJKLMNOPQRSTUVWXY

Zjednodušení

Když je klíč delší, na každé pozici se šifruje jiným bajtem klíče. To je dobré pro zakrytí otevřeného textu, ale špatné pro „driver“, kterému se musí říci, na jaké pozici je daný znak. Kdyby všechny znaky klíče byly stejné (tj. D = 1), šifrovali bychom na každé pozici stejně, a to tak, že znak d bychom zaměnili za znak d xor k a nepotřebovali bychom udávat pozici znaku. Používali bychom tedy jednu substituci jako Caesar (*tabulka 1*), další z metod, kterou nám Příroda někam šikovně zakódovala.

Polyalfabetická substituce

Zvolíme-li kvalitnější klíč o více znacích (D >1), máme k šifrování více šifrových abeced, a to na každou z D pozic jednu. Naši předchůdci, kteří neměli počítače, používali abecedu o 26 znacích a místo binární operace xor používali přičítání, přesněji posun v abecedě (když byl klíč A, posun byl o nula, B o jednu pozici atd.). Dnes bychom jednoduše řekli přičítání modulo 26, takže operace by měla tvar (d + k) mod 26, ale dříve se tato operace vyjadřovala tabulkou, latinsky Tabula recta (*tabulka 2*) (dále jen „Tabula“). V záhlaví Tabuly je abeceda otevřená, v prvním řádku je první šifrová abeceda (označme ji pracovně A-abeceda), v druhém řádku je druhá šifrová abeceda atd., až v 26. řádku je Z-abeceda. Na této tabulce si ukážeme vývoj šifrování.

Jak vznikla Vigenerova šifra

Caesar (50 let př. n. l.) proslul svojí šifrou, kdy písmena posouval o tři místa v abecedě, tedy používal D-řádek. Později se začala používat nikoli posunutá abeceda, ale zpřeházená (zamíchaný řádek). V roce 1400 je už známa frekvenční analýza a luštění takové substituce. Pro posun se používaly i mechanické pomůcky, jednou z nich byl šifrovací disk, který vynalezl L. Alberti v roce 1467. Sestával ze dvou koleček, na jejichž obvodu byla otevřená a šifrová abeceda. Jakýkoliv X-řádek Tabuly se dal vytvořit nastavením vnějšího písmene X nad vnitřní písmeno A. Tato šifrovací pomůcka byla jednoduchá a používala se 500 let na (*obr. 1* je v americké verzi z doby občanské války). Alberti však poznamenal, že během šifrování zprávy by se mohlo nastavení kotoučů měnit, čímž objevil polyalfabetickou substituci. V roce 1518 J. Trithemius vydal první tištěnou kryptologickou knihu, Polygraphiae, která obsahuje *tabulku 2*. Navrhoval, aby se šifrovalo postupně první písmeno A-abecedou, druhé B-abecedou atd. „Náš“ Vigenerův systém vynalezl G. Belaso v roce 1553 a popsal ho v knize La Cifra. Zvolil tajný krátký klíč (například KLIC), pomocí něhož vybíral řádky Tabuly, kterými šifroval: K-řádek, L-řádek, I-řádek, C-řádek a poté od K-řádku zase dokola. G. Porta v roce 1563 v knize De Furvitis Literarum Notis popsal obecný princip polyafabetické substituce. Navrhl, aby Tabula recta obsahovala zpřeházené abecedy a Belasův klíč byl co nejdelší. Konečně B. Vigenere v roce 1585 v knize Pojednání o šifrách popsal tzv. autoklíč, ale z nějakého důvodu se právě Belasův systém začal nazývat Vigenerovým (?!). Zůstal 300 let neluštitelným! Jeho rozluštění popsal až v roce 1863 F. Kasiski. V roce 1917 G. Vernam použil Vigenerův systém s binární abecedou a prodloužil délku Belasova klíče do nekonečna. Tím vznikla teoreticky i prakticky nerozluštitelná šifra za předpokladu, že klíč je náhodný a je použit jen k šifrování jedné zprávy. Důkaz, že tuto šifru nelze rozluštit, podal až v roce 1949 C. E. Shannon. Je to jedna z mála šifer, kde nezáleží na tom, jak je útočník chytrý a kolik má počítačů, včetně kvantových. Zprávu bez klíče nikdy nerozluští. V tom je velký rozdíl od Vigenerovy šifry, jejíž luštění je školní úlohou.

Obr. 1 Šifrovací disk

A propos, kde dnešní programátoři berou klíč k oné Vigenerově šifře, kterou používají? Vrátíme-li se k naší aplikaci, bude muset klíč asi zadat její uživatel. Pak se ale zjistí, že je to neprůchodná komplikace, takže nakonec program použije nějakou „tajnou konstantu“. O kvalitě takového šifrování nelze samozřejmě ani hovořit, nicméně původní účel to splňuje.

Vlastimil Klíma, Tomáš Rosa,
v.klima@volny.cz, tomas.rosa@rb.cz

LITERATURA

[1] E-archivy http://cryptography.hyperlink.cz, http://crypto.hyperlink.cz