Složitost v kryptologii

V předchozím čísle ST jsme se zabývali tím, jak narůstá složitost jednotlivých bitů součtu (ADD) dvou 32bitových čísel s = a + b. Viděli jsme, že bity s_i jakožto funkce všech bitů a_j a b_j (j = 0, ..., i) jsou velmi složité funkce. Bude-li složitost definována jako počet členů booleovské funkce při jejím zápisu v algebraické normální formě, složitost bitů s_i narůstá exponenciálně. Bylo uvedeno, že funkce pro nejvyšší bit součtu s_{31} obsahuje neuvěřitelný počet 2 147 483 649 termů.

Složitost v elektronice

Jiná míra složitosti je velmi blízká obvodové realizaci funkcí. Kryptologům nejvíce vadí operace AND (*), zatímco XOR (⊕) nikoliv, takže jedna z možných definic složitosti booleovské funkce je dána nejmenším počtem operací AND nutných k její realizaci. Přitom počet operací XOR ani počet meziproměnných, které při výpočtu vznikají (paměťové buňky pro bitové mezivýsledky), se nepočítají, přestože to může být doplňková míra. To odpovídá počtu operací AND v obvodu, který realizuje danou funkci.

Složitost operace ADD

Spočítejme si složitost operace ADD. Při sčítání vznikají mezivýsledky jako bity přenosu c_1, c_2, ..., c_{31} a pro bity součtu jsou vztahy

$s_i = a_i \oplus b_i \oplus c_i$, i = 0, ..., 31, kde $c_0 = 0$

a ostatní bity přenosu jsou dány funkcemi

$c_1 = a_0\, b_0$

$c_2 = a_1\, b_1 \oplus a_1\, c_1 \oplus b_1\, c_1$

$c_3 = a_2\, b_2 \oplus a_2\, c_2 \oplus b_2\, c_2$

$c_4 = a_3\, b_3 \oplus a_3\, c_3 \oplus b_3\, c_3$

…

$c_{30} = a_{29}\, b_{29} \oplus a_{29}\, c_{29} \oplus b_{29}\, c_{29}$

$c_{31} = a_{30}\, b_{30} \oplus a_{30}\, c_{30} \oplus b_{30}\, c_{30}$

Složitost bitů součtu se ze čtyř miliard rapidně snížila na neuvěřitelných 30 × 3 + 1 = 91 operací AND. Při optimalizaci lze operaci $c_{i+1} = a_i\, b_i \oplus a_i\, c_i \oplus b_i\, c_i$ zapsat s použitím pouze dvou operací AND jako $c_{i+1} = a_i\, (b_i \oplus c_i) \oplus b_i\, c_i$. Tím se dosáhne složitosti 61 operací AND. Optimalizaci však lze ještě vylepšit! Když se c_{i+1} vyjádří jako $(b_i \oplus c_i)\, (a_i \oplus c_i) \oplus c_i$, docílí se skutečného minima a na realizaci celého součtu a + b bude třeba pouze 31 operací AND. To je dobré, je dosaženo minima. Každý člen c_1, ..., c_{31} totiž obsahuje nelineární člen, čili musí použít alespoň jednu operaci AND. ADD má tedy přesně složitost 31.

Porovnávání šifer

Uvedenou definici je možné použít k vyjádření složitosti např. blokových šifer, hašovacích funkcí nebo jakýchkoliv jiných kryptografických technik. Žádná míra není dokonalá a ani zde navržená míra není absolutně spravedlivá pro všechny funkce. Proto se obvykle porovnávají výsledky pro více měr nebo více vlastností. Ale toto je velmi jednoduchá a velmi účinná míra a zbývá ji jen dobře použít. Například u blokových šifer, které mají různou délku bloku, je vhodné vypočítat počet operací AND pro zpracování celého bloku a poté složitost vztáhnout k jednomu bitu. Tím lze porovnávat složitost blokových šifer s různou délkou bloku, různou délkou klíče apod. Také je vhodné měřit složitost za předpokladu, že klíč je konstanta, a za předpokladu, že klíč je proměnná (když se luští otevřený text), a za předpokladu, že otevřený text je konstanta (luští se klíč) nebo že je to také proměnná, takže hned je zřejmé, že to jsou tři různé hodnoty složitosti. Počet operací AND se u těchto variant může lišit, protože operace AND s konstantou nemá žádnou složitost.

Tabulka 1 Nejrychlejší kandidáti na SHA-3 a jejich utajená míra složitosti

Algoritmus	Rychlost (cyklů/bajt)	Počet operací AND na jeden bit zprávy
SHA-1	9	17
BMW	7	24
BLAKE	9	29
Shabal	10	13
CubeHash	13	992
SIMD	12	23
Skein	21	26
SHA-2	20	40

AND a ADD rozhodují o bezpečnosti!

Zdálo by se, že je zde věnován prostor nezáživné teorii, ale není tomu tak. Na počtu operací AND skutečně velmi závisí bezpečnost blokových šifer a jiných kryptografických technik! Příslušné problémy hledání klíče, kolizí apod. lze vždy přepsat do logických proměnných a booleovských rovnic. Jestliže nejsou složité, poradí si s nimi programy, které se nazývají SAT solvery. Jde o vědeckou oblast, která se rychle rozvíjí a neustále vylepšuje své výsledky. Není-li nelinearita použité funkce dostatečná, SAT solver by ji prolomil. Proto počty operací AND skutečně rozhodují o bezpečnosti. Pro ilustraci jsou v *tabulce 1* uvedeny nejrychlejší z patnácti kandidátů na novou normu pro hašování SHA-3, z nichž bude letos v srpnu až září vybráno pět nejlepších jak z hlediska rychlosti, tak bezpečnosti.

Tajná míra složitosti?

Poznamenejme, že v tabulce byly uvažovány varianty algoritmů s 256bitovou haší a výpočet probíhal na 32bitových PC. Nezávisle na rychlosti procesoru je zde uvedena rychlost v cyklech na bajt. Pro vysvětlení: např. s PC s taktem 7 GHz (aby se to dobře počítalo), tj. se sedmi (giga)cykly za sekundu, se bude pomocí BMW hašovat rychlostí jeden (giga)bajt za sekundu, tj. 1 GB/s. V *tabulce 1* je zřejmá neuvěřitelná shoda na složitosti podle zde uvedené míry. Je nutné mít na zřeteli, že tyto algoritmy jsou velmi rozdílné, že vznikaly nezávisle a v tajnosti a do jejich návrhu hovořili tři až čtrnáct vědeckých pracovníků u každého algoritmu. Jak je možné, že ty nejrychlejší algoritmy mají velmi nápadně stejný počet operací AND? Jejich popisy jsou velmi obtížně srovnatelné, a přitom jako by zde existovalo určité tajné pravidlo v pozadí, které tyto algoritmy nevědomky dodržují. Toto pravidlo lze nazvat „chytrá složitost“, neboť tuto myšlenku mají algoritmy společnou – co nejúčinněji využít stávajících rychlých, a přitom co nejsložitějších operací k získání kryptograficky silné funkce. To, že výsledky mohly být velmi různorodé, ukazuje příklad jednoho z kandidátů, který zcela vybočuje. Vyvolává to mnoho otázek, na něž u CubeHash odborníci zatím nedokážou odpovědět. Jedno je ale jisté, že operace AND tady hraje důležitou roli ve složitosti, a to takovou, že se na ní nevědomky shodne pět nezávislých týmů. Dva algoritmy přitom používají velmi různorodé funkce (modulární násobení nebo přímo AND), zatímco čtyři zbylé používají jako jedinou nelineární pouze operaci ADD, jejímž prostřednictvím do obvodů vnášejí 31 operací AND.

Závěr

Pokud je kryptografická technika udělána dobře, je řešení soustavy booleovských rovnic vyjadřujících obvodové schéma jediná cesta, jak tuto techniku prolomit. Proto se zdá, že zde navržená míra složitosti je přesně to, co je zapotřebí k ohodnocení bezpečnosti, a kromě toho je to míra velmi jednoduchá, ověřená a účinná.

Vlastimil Klíma,
vlastimil.klima@knzsro.cz

LITERATURA

[1] Archiv autora a článků kryptologie pro praxi: http://cryptography.hyperlink.cz.