Posudek oponenta na diplomovou práci Bc. Daniela Joščáka
Hledání kolizí v hašovacích funkcích.

Práce psaná v angličtině se zabývá nedávno zveřejněnými útoky na hašovací funkci MD5. Šest kapitol lze obsahově rozdělit na tři základní části:

- popis hašovací funkce a princip útoku
- vlastní algoritmus pro hledání kolizí
- analýza složitosti a srovnání s konkurenčními algoritmy.

Práce vychází ze znalosti zcela nedávných publikací a předpokládá jejich dobré zvládnutí.

Z teoretického hlediska spoléhá autor na dále neanalyzovaný poznatek z literatury, podle nějž je kolizní dvojice konstruována tak, aby splňovala danou diferenční cestu. Samotné toto východisko je zcela legitimní, je ale škoda, že je ponecháno do velké míry na čtenáři, aby si tuto výchozí situaci uvědomil. Např. vztah mezi R_{19} a R'_{19} na s. 18 tak působí poněkud záhadně.

Principem algoritmu je omezit počet prohledávaných zpráv (resp. zvýšit pravděpodobnost nalezení kolize) pomocí vhodně zvolených vynucených informací. Detailní popis algoritmu přesahuje možnosti práce, autor proto čtenáři vysvětlí pouze jeho základní principy. V porovnání s komplikovaností algoritmu je popis poměrně srozumitelný. Opět ovšem chybí výslovná analýza toho, proč právě zvolené parametry se deterministicky dopočítávají a jiné se volí náhodně.

Základním úspěchem práce je jistě funkčnost algoritmu, která je srovnatelná s nejlepšími existujícími konkurenty (v tomto vycházím z tvrzení uvedených v práci). Důležitým faktem je také to, že algoritmus autora byl vytvořen nezávisle na zcela nedávno publikovaných algoritmech Stevense a Klímy. Na druhou stranu by diskuse o důvodech zvoleného postupu a možnostech jeho modifikace, nebo alespoň uvedení důvodů pro absenci této diskuse, práci prospěly. Proč např. není vhodné dopočítávat deterministicky kandidáty, kteří splňují podmínky pro registry R_j, $j > 18$?

V kapitole 5 srovnává autor svůj algoritmus s algoritmy Stevense a Klímy. Z teoretického hlediska je zajímavá snaha o odhad složitosti algoritmu. Základním předpokladem správnosti tohoto odhadu je tvrzení o vzájemné nezávislosti podmínek na jednotlivé registry a předpoklad rovnoměrného rozdělení pravděpodobnosti jejich splnění. Tyto dva předpoklady bohužel nejsou nijak diskutovány, a dosažený odhad je proto možno považovat pouze za heuristický.

Práce obsahuje řadu drobných překlepů. Z těm vážnějším patří:

- s. 37, rovnost (22): místo $IV^0_{1,25}$ má být $IV^0_3[25]$
- s. 38: je zmaten vztah mezi M_i a W_i
- Proposition 5.3.: v druhé polovině tvrzení chybí význam uvedených čísel.

Celkové hodnocení. : Autor prokázal schopnost pracovat s literaturou i schopnost vlastního přístupu. Práce splňuje požadavky na diplomovou práci.

Hlavním nedostatkem práce je nedostatečně formulovaný teoretický rámec vytvořeného algoritmu.

Navrhuji známku *velmi dobře*.

Praha 22. května 2006

Mgr. Štěpán Holub, Ph.D.