MD5 Cracking

Vložil/a Stoyan [1], 5 Říjen, 2007 - 11:07

- Cracking [2]
- Security [3]

Hashovací funkce jsou jednosměrné. To znamená, že z řetězce vytvoříme hash jednoduše, ale pokud známe pouze hash, mělo by být odhalení původního řetezce (hesla) nemožné. U MD5 je to pravda pouze z části...

V případě úspěšného proniknutí do webové aplikace / databáze, se můžeme často setkat s hesly v tomto tvaru: **02fec283a4a59afa3678c49d09cc7805**. Nejedná se o vlastní heslo, ale o tzv. otisk hesla vytvořený pomocí hashovací funkce MD5 [4]. Hashovací funkce jsou jednosměrné. To znamená, že z řetězce vytvoříme hash jednoduše, ale pokud známe pouze hash, mělo by být odhalení původního řetezce (hesla) nemožné. U MD5 je to pravda pouze z části.

Od srpna roku 2004 je znám postup k nalezení kolize dvou řetězců. V praxi to vypadá tak, že vám program vyplivne dva řetězce, které mají stejný hash (to by se hashovacím funkcím stávat nemělo). To je ale nám, kteří chceme z hashe dostat původní řetězec, na nic. Tyto vygenerované řetězce jsou velice dlouhé a často obsahují i netisknutelné znaky. Tato metoda trvala dříve asi 8 hodin, ale díky objevu předního českého kryptoanalytika Vlastimila Klímy se tato doba snížila na cca jednu minutu. Zde je homepage jeho projektu [5]. V případě, že tedy někde uslyšíte, že "MD5 byla prolomena" nemluví se o ničem jiném, než o kolizi dvou řetězců. Zjistit ale z hashe původní řetězec pomocí určitého algoritmu, je však nadále nemožné.

My ale máme hash a potřebujeme zjistit původní řetězec. Jaké tedy máme možnosti? V podstatě jsou asi čtyři. A pokud tyto selžou, není prakticky šance, že původní řetězec naleznete. Podívejme se tedy podrobněji na jednotlivé metody.

Brute-force attack

Tento způsob luštění je založen na testování všech možných řetězců o zvolené délce nad zvolenou znakovou sadou. Tou může být velká a malá abeceda, číslice a speciální znaky. Dejme tomu, že si zvolíme, že chceme generovat řetězce o délce 1 - 5 a v nich používat pouze malá písmena anglické abecedy. Těch je 26. Celkový počet vygenerovaných řetězců tedy bude 26^1 + 26^2 + 26^3 + 26^4 + 26^5 + 26^6. Vyjde nám obrovské číslo. Z každého vygenerovaného řetězce je vytvořen hash a ten je porovnán s naším hashem. Pokud původní řetězec (heslo), který hledáme, byl dlouhý 1 až 5 znaků a obsahoval pouze malá písmena anglické abecedy, program ho během několika sekund odhalí. Tímto způsobem pracuje brute-force attack. Vypadá to ideálně, ale praxe není tak růžová. Hesla jsou většinou delší než pouhých pět znaků a obsahují i velká písmena, číslice a znaky. Vygenerování a porovnání tak obrovského množství řetězců trvá celkem dlouhou dobu. Zkuste si vypočítat počet kombinací, který je potřeba na prolomení devítimístného hesla, složeného z čísel a velkých a malých písmen. Dostanete tak obrovské číslo, že by vygenerování takového počet kombinací řetězců zabralo klidně i několik desítek let. Tato metoda je v praxi tedy využitelná na hesla tak do 6 až 7 znaků - záleží na tom, jak dlouho jste ochotni čekat. Programem, který zvládne louskat MD5 (brute-force i dictionary(viz dále)) je například Cain & Abel [6].

Pro představu jsem sestavil tabulku, ze které je zřejmá časová náročnost louskání MD5 hashů různě dlouhých hesel programem Cain & Abel (předpokládal jsem, že hesla obsahují pouze malá písmena anglické abecedy a číslice):

HesloZnakůHashČasa2b347cf26c17f2ef17fb3d3a5967b83dbbd

zlomek vteřinyxs3o4689101f13c29393b4cfa696662433e3fzlomek
vteřinyd9u2k5beb6bb3b5cf6abdc9c915c71fc3d52906
sak74ka69b9b12c2fe191e341d8a88159c6f8d3326
sheslo1271d7efc5be9a522837bd7880072dd00423
hheslo12386a284155906c26cbca20c53376bc63ac> 10 h

Dictionary attack

Neboli také slovníkový útok je metoda podobná té předchozí až na to, že řetězce nejsou generovány, ale jsou brány z externího souboru. Ten se nazývá slovník a většinou to je obyčejný textový soubor, který na každém řádku obsahuje jedno slovo. Při tomto typu louskání záleží hlavně na slovníku jaký použijete a také hodně na štěstí.

Rainbow tables

Rainbow tables jsou tabulky s předgenerovanými hashi. To znamená, že hashe se nemusí generovat z náhodně zkoušených řetězců, ale rovnou se vyhledávají v tabulce. Velikost tabulek se pohybuje od 700MB (3-14 znaků: malá písmena) až třeba po 40GB (1-14 znaků latin1) a každý si může předgerovat svoji tabulku přesně na míru. Může si vybrat rozsah, typ hashe (md5) a znakovou sadu. Prolomení hashe se pak pohybuje kolem 10 minut. Tak to byla teorie. Já Rainbow tables v praxi nezkoušel, hlavně kvůli nedostaku místa na disku. Tabulka, která by za něco stála by totiž měla asi 64 GB, což by bylo ještě snesitelné, jenže na Pentiu 3 GHz by se generovala asi tři týdny, což si v reálu nedovedu představit. Pokud někdo s Rainbow tables máte zkušenosti, uvítám je v komentářích. Ještě přidám odkaz na Project RainbowCrack [7], kde se dozvíte více.

MD5 Online cracking

Toto je vlastně spojení rainbow tables a databáze hashů, ke kterým ze známý původní řetězec. Vše navíc funguje přes webový prohlížeč jako webová aplikace, do které pouze vložíte hash a aplikace vám vyplivne původní řetězec. No nezní to krásně? Samozřejmě, že to nefunguje vždy. Tipuji, že například takový hash pro řetězec A.x4a&/s4-s5 v online databázi nenajdete, ale zase na druhou strany si přiznejme, jak dlouho by trvalo jeho odhalení pomocí brute-force. Webů, zaměřujících se na tuto "službu" je více. Zde je několik příkladů:

http://passcracking.ru [8]
http://www.milw0rm.com/cracker/insert.php [9]
http://www.gdataonline.com/seekhash.php [10]
http://md5.rednoize.com/ [11]
http://plain-text.info/ [12]

Nyní si pojďme shrnout, jaké máme tedy možnosti v případě, že získáme hash a potřebujeme z něj zjistit původní řetězec. Nejprve se určitě vyplatí navštívit několik webů s online MD5 crackery (viz linky nahoře). Nestojí to téměř žádný čas a v případě, že se daný hash už nachází v databázi, dostanete okamžitě původní řetězec. V případě, že zde s hashem neuspějete, nezbývá vám nic jiného než brute-force /dictionary attack.

Z tohoto článku vyplývá také jisté ponaučení. Že nevíte jaké? Z tabulky nahoře jasně vyplývá závislost délky hesla a použitých znaků na době louskání. Jako minimální délku hesla bych zvolil 8 znaků. Pokud tedy použijete heslo například 'heslo1' a útočník se dostane k jeho MD5 hashi, nebude pro něj velký problém heslo zjistit. Naproti tomu hash takovéhoto hesla 'h0Es0_k_M4ilU' bude pro útočníka neprolomitelný.

URL článku: https://security-portal.cz/clanky/md5-cracking

Odkazy:
[1] https://security-portal.cz/users/stoyan

[2] https://security-portal.cz/category/tagy/cracking
[3] https://security-portal.cz/category/tagy/security
[4] http://en.wikipedia.org/wiki/MD5
[5] http://cryptography.hyperlink.cz/MD5_collisions.html
[6] http://www.oxid.it/cain.html
[7] http://www.antsight.com/zsl/rainbowcrack/
[8] http://passcracking.ru
[9] http://www.milw0rm.com/cracker/insert.php
[10] http://www.gdataonline.com/seekhash.php
[11] http://md5.rednoize.com/
[12] http://plain-text.info/