Szyfrowanie, prosta sprawa

Tomasz Marcinek

Opracowana przez młodego wrocławianina funkcja jednokierunkowa VMPC pozwala budować bardzo wydajne szyfry strumieniowe.

Współczesna kryptografia matematyką stoi. Jeden z jej działów okazał się bowiem szczególnie przydatny do zabezpieczania informacji przed niepowołanym wzrokiem ciekawskich. Wyjątkowość tzw. funkcji jednokierunkowych polega na tym, że najprościej rzecz ujmując nie da się lub szalenie trudno jest odgadnąć jej argument(y) (wartości wejściowe) na podstawie wyniku przekształcenia.

Funkcji jednokierunkowych używamy wszyscy na co dzień. Logując się do systemów Unix/Linux, nawet nie zdajemy sobie sprawy, że nasz login i hasło są przekształcane przez funkcję jednokierunkową, np. MD5, by nikt nie mógł ich podsłuchać. Łącząc się za pomocą przeglądarki internetowej z witryną banku, wykorzystujemy funkcję jednokierunkową zawartą w protokole SSL. Funkcji jednokierunkowych (różnych) używa się także podczas składania podpisu elektronicznego.

Przekształcenia danych wykonywane przez powszechnie stosowane funkcje jednokierunkowe (popularnie nazywane haszującymi) charakteryzują się zwykle dużym stopniem skomplikowania, powtarzając wielokrotnie wiele stosunkowo skomplikowanych operacji, by uniemożliwić zdekodowanie oryginalnej treści. Skomplikowanie musi mieć jednak granice, ponieważ większość operacji szyfrowania jest wykonywana programowo. Skomplikowanie zwiększa ponadto ryzyko popełnienia błędu w konstrukcji algorytmu. Od lat matematycy pracują więc nad takimi funkcjami jednokierunkowymi i algorytmami kryptograficznymi, które są zarówno proste (a więc i szybkie), jak i bezpieczne.

Nieoczekiwanie pomysł na prostą funkcję jednokierunkową i oparty na niej prosty, a jednocześnie bezpieczny szyfr strumieniowy zrodził się w Polsce. Jego autorem jest Bartosz Żółtak, niedawny absolwent... Marketingu i Zarządzania na Politechnice Wrocławskiej (wywiad z nim zamieszczamy na str. 22). Opracowana przez niego funkcja jednokierunkowa VMPC składa się jedynie z trzech podstawowych operacji, co przekłada się na bardzo dużą wydajność szyfrowania. Na podstawie tej funkcji Bartosz Żółtak zbudował strumieniowy algorytm szyfrujący działający w przybliżeniu dwukrotnie szybciej niż algorytm AES.

Jak konstrukcja cepa

Nowa funkcja jednokierunkowa opiera się na bardzo prostym przekształceniu. Tak prostym, że wielu specjalistów nie chciało uwierzyć, iż w ogóle jest to możliwe. Posłużmy się prostym przykładem. Pierwsza linia tabeli 1 zawiera numer kolumny (indeks). Poniżej umieszczono dowolnie wylosowany ciąg liczb (bajtów) z zakresu od 0 do 5 - to ciąg początkowy, a dokładnie permutacja ze zbioru (0, 5). W pierwszym kroku następuje przypisanie każdej z otrzymanych wartości permutacji numerowi indeksu. Dla elementu o indeksie 0 wartość permutacji wyniosła 2, dla indeksu 1 - 5 itd.

W drugim kroku sprawdzamy, jaką wartość przyjmuje każdy element permutacji P dla indeksu równego jego wartości. Poczynając od lewej strony, wartości 2 jako indeksowi odpowiada wartość permutacji P = 4, zaś wartości 5 traktowanej jako indeks - P = 1 itd. W trzecim kroku powtarzamy czynność podstawiania, traktując jako ciąg wyjściowy ciąg wartości otrzymanych w kroku drugim. Kluczowa różnica polega na tym, że do wyniku (wartości, która przechodzi na indeks) dodajemy 1. Tam, gdzie wynik przekroczy dopuszczalną wartość z zakresu - w tym przypadku 5 - stosuje się operację modulo

("przekręcenie licznika", w wyniku którego wartość 6 oznacza 0, 7 oznacza 1 itd.). W ten sposób otrzymujemy kolejne wartości funkcji Q = $VMPC_1(P)$ pierwszego stopnia.

Zbyt proste? Można to skomplikować. Aby otrzymać wartości funkcji VMPC drugiego stopnia, wystarczy w kolejnym (IV) kroku wartości uzyskane potraktować jako permutację początkową, podstawić jej wartości za indeksy - lecz dodając 2 i dla każdego z nich odczytać wartość permutacji P. Tak powstaje wartość funkcji VMPC drugiego stopnia - Q = $VMPC_2$ (P). Dodanie wartości zaburza strukturę cykli permutacji P, a więc strukturę zależności między indeksami a wartościami permutacji P. Dodawanie w każdej kolejnym przekształceniu innej wartości (zamiast 2 powyżej można było podstawić dowolną liczbę inną niż 1) dodatkowo podnosi trudność odgadnięcia początkowej permutacji (2, 5, 4, 0, 3, 1). Właściwość tę oddaje nazwa funkcji. VMPC to skrót od Variably Modified Permutation Composition - zmiennie modyfikowane złożenie permutacji.

Nie do złamania

Opisana funkcja, mimo swojej prostoty, okazuje się niezwykle skuteczną metodą kodowania. Oczywiście, zbiór sześcioelementowy nie stanowi wyzwania dla dzisiejszych mocy obliczeniowych. Gdy jednak wartości jest 256, sprawy mają się zupełnie inaczej. Na podstawie badań empirycznych autor funkcji twierdzi, że do odgadnięcia oryginalnej permutacji składającej się z 256 elementów potraktowanej funkcją VMPC pierwszego stopnia konieczne jest odgadnięcie co najmniej 34 jej wartości (nie muszą to być liczby kolejne).

Stosując skomplikowany algorytm wnioskowania, jest możliwe odtworzenie pozostałych elementów P. Próba zrealizowania tego zadania przekracza jednak jakiekolwiek istniejące bądź przewidywalne moce obliczeniowe. Do tego celu jest bowiem konieczne sprawdzenie średnio aż 2^{260} kombinacji. Jest to więcej niż kwadrat uznanej dziś za standard mocy kryptograficznej 2^{128}.

Na bazie funkcji jednokierunkowej VMPC zbudowano szyfr strumieniowy o tej samej nazwie.

To obecnie najprostszy publicznie znany bezpieczny algorytm szyfrowania danych. Do realizacji szyfrowania wykorzystuje on opisaną wcześniej funkcję VMPC oraz dwie dodatkowe operacje, których celem jest zapewnienie jeszcze wyższej odporności na złamanie niż w przypadku samej funkcji (moc kryptograficzna ok. 2^{900} operacji) oraz zapewnienie, że generowany przez szyfr strumień danych dowolnej długości nie da się praktycznie odróżnić od ciągu losowego. Pod tym względem szyfr VMPC znacznie przewyższa właściwościami najpopularniejszy i najpowszechniej stosowany szybki szyfr strumieniowy - RC4.

Więcej informacji pod adresem: http://www.vmpcfunction.com

Jak działa szyfr strumieniowy?
Szyfr strumieniowy to przekształcenie matematyczne pozwalające na podstawie dowolnego początkowego ciągu znaków (bajtów) uzyskać inny ciąg znaków o takich właściwościach, że na podstawie ciągu wynikowego nie jest w praktyce możliwe odgadnięcie sekwencji ciągu początkowego. Druga ważna właściwość szyfrów strumieniowych polega na tym, że wartości ciągu wynikowego nie powinny się różnić od ciągu losowego, a więc nie wykazywać żadnych regularności umożliwiających odgadnięcie ciągu pierwotnego za pomocą analizy statystycznej.

W zastosowaniach kryptograficznych, a dokładnie w kryptografii symetrycznej, ciąg początkowy to znany obu komunikującym się stronom klucz szyfrujący. Szyfr strumieniowy generuje na jego podstawie (pseudo)losowy ciąg bajtów, które są następnie "mieszane" z

danymi mającymi podlegać szyfrowaniu. Mieszanie polega na dodawaniu do siebie wartości, bajt po bajcie lub (w praktyce najczęściej) wykonywaniu na parze: bajt (pseudo)losowy i bajt danych operacji bitowej XOR. Po drugiej stronie łącza następuje dekodowanie - znając klucz, można na jego podstawie wygenerować ten sam ciąg (pseudo)losowy i od otrzymanych od nadawcy wartości odejmować wartości ciągu lub poddawać napływające wartości i samodzielnie wygenerowany ciąg (pseudo)losowy operacji XOR. Ciąg otrzymanych w ten sposób wartości będzie zawierać te same dane, które nadawca zaszyfrował.

Podobną do klucza, ale dodatkową rolę pełni tzw. wektor inicjujący, będący drugim - obok klucza szyfrującego - parametrem algorytmu szyfrującego. Różnica między nimi polega na tym, że klucz jest generowany/zmieniany rzadziej niż wektor inicjujący, m.in. w przypadku szyfrowania wiadomości e-mail każda kolejna wiadomość jest szyfrowana z użyciem innego wektora. Pozwala to uzyskać tę własność, że szyfr nawet dla wielokrotnie użytego tego samego klucza - ale z innym dla każdej wiadomości wektorem inicjującym (który przesyłany jest jawnie) - wygeneruje za każdym razem inny strumień danych, co pozwala bezpiecznie szyfrować praktycznie dowolną liczbę wiadomości przy użyciu tego samego klucza.

Strumieniem albo w blokach

Algorytmy kryptograficzne dzielą się zasadniczo na blokowe i strumieniowe. W pierwszym przypadku szyfrowaniu podlegają porcje danych (zwykle 8 lub 16 bajtów), w drugim zaś pojedyncze bajty.

Algorytmy blokowe to konstrukcje sprawdzone i powszechnie używane (np. algorytm DES). Ich mankamentem jest słaba wydajność, objawiająca się dużymi opóźnieniami w transmisji (np. w IPsec). Wynika to nie tylko z faktu, że składają się one z wielu skomplikowanych etapów obliczeniowych, ale także dlatego że do ich bezpiecznej pracy są potrzebne dodatkowe operacje określające tryb łączenia bloków (np. CBC czy OFB).

Algorytmy strumieniowe są, historycznie rzecz biorąc, młodsze. Oprócz braku etapu łączenia bloków stosuje się w nich także efektywniejsze funkcje jednokierunkowe, w rezultacie czego są one wydajniejsze niż szyfry blokowe.