Podpis elektroniczny – zasady działania

dr inż. WOJCIECH NOWAKOWSKI, mgr inż. ROBERT POZNAŃSKI

Instytut Maszyn Matematycznych, Warszawa

Przez bardzo długi czas techniki kryptograficzne wykorzystywane były do szyfrowania i utajniania wiadomości i informacji. Praktyki takie stosowano już za czasów Cesarstwa Rzymskiego za pomocą prostego algorytmu podstawieniowego (tzw. szyfr cezara). Każda litera szyfrogramu była zastępowana inną, odległą o kilka pozycji w alfabecie (a -> d, b-> e itd.). W tym przypadku kluczem do odszyfrowania była wiedza w jaki sposób przestawione są litery alfabetu. Szczytowym osiągnięciem kryptografii sprzed ery komputerów i maszyn liczących była niemiecka maszyna szyfrująca Enigma. Miała ona budowę mechaniczno–elektryczną. Najważniejszymi elementami były: 26-znakowa klawiatura oraz zespół kilku, najczęściej 3-5 rotorów oraz jednego nieruchomego rotoru odwracającego. Każdy z nich posiadał 26 styków odpowiadających kolejnym literom alfabetu. Często między klawiaturą a zespołem wirników, znajdowała się centralka pozwalająca na ręczną zmianę znaków przy pomocy kabelków łączących poszczególne litery. Po naciśnięciu klawisza obwód elektryczny zamykał się, a prąd przepływał przez elementy składowe maszyny ostatecznie powodując zapalenie się jednej z wielu lampek podświetlających literę wyjściową. Ciągłe obroty wirników zmieniały drogę jaką przebywał sygnał. Enigma szyfrowała tzw. szyfrem poliaflabetycznym.

Jednym z pierwszych zaawansowanych algorytmów szyfrujących jest zatwierdzony przez NIST w 1976 roku *Data Encryption Standard*, znany szerzej jako **DES**. W tym algorytmie, jak i wszystkich poprzednich do szyfrowania i deszyfrowania używano tego samego klucza. Były to tzw. algorytmy symetryczne. W tym samym 1976 roku Martin Hellman i Whitfield Diffi opublikowali nowy pomysł, który polegał na zastosowaniu dwóch powiązanych ze sobą matematycznie kluczy: publicznego oraz prywatnego, z których jeden służy do szyfrowania, a drugi do deszyfrowania wiadomości. Tak powstała kryptografia asymetryczna, która stała się podstawą podpisu elektronicznego.

Podpis cyfrowy a elektroniczny

Pojęcie ***podpisu cyfrowego*** (*digital signature*) zostało zdefiniowane przez normę ISO 7498-2:1989 jako „dane dołączone do danych lub ich przekształcenie kryptograficzne, które pozwala odbiorcy danych udowodnić pochodzenie danych i zabezpieczyć je przed fałszerstwem”. Podpis cyfrowy jest pojęciem szerszym niż podpis elektroniczny. Podpis cyfrowy nie musi być generowany przez człowieka, do tej kategorii zalicza się np. zastosowania matematycznej operacji „podpisywania cyfrowego” wykorzystywane np. w protokołach kryptograficznych, które „podpisują” np. tymczasowe liczby losowe w celu potwierdzenia posiadania klucza prywatnego. Pojęcie ***podpis elektroniczny*** (*electronic signature*) jest natomiast wprowadzone przez unijną Dyrektywę 1999/93/EC i jednoznacznie określa, że jest to operacja podpisywania konkretnych danych (dokumentu) przez osobę fizyczną. Podpis elektroniczny to w istocie dodatkowa informacja dołączona do wiadomości, służąca do weryfikacji jej źródła. Norma PN-I-02000 podaje definicję: podpis cyfrowy „to przekształcenie kryptograficzne danych umożliwiające odbiorcy danych sprawdzenie autentyczności i integralności danych oraz zapewniające nadawcy ochronę przed sfałszowaniem danych przez odbiorcę”. Cztery główne warunki jakie muszą być spełnione przez to przekształcenie:

– uniemożliwienie podszywania się innych pod daną osobę (uwierzytelnienie osoby, autentyfikacja),
– zapewnienie wykrywalności wszelkiej zmiany w danych transakcji (integralność transakcji),
– zapewnienie niemożliwości wyparcia się podpisu przez autora,
– umożliwienie weryfikacji podpisu przez osobę niezależną.

RSA

Podstawową cechą uniemożliwiającą zastosowanie w procedurze podpisu elektronicznego szyfrowania symetrycznego (z jednym kluczem) jest to, że klucz musi być przekazany odbiorcy przez nadawcę informacji. Po co szyfrować wiadomość, jeżeli klucz do jej odszyfrowania może być przejęty przez osoby trzecie? Dlatego do szyfrowania wykorzystuje się algorytmy symetryczne z dwoma kluczami: publicznym, jawnym i zależnym od niego kluczem prywatnym. Pierwszy z nich służy do szyfrowania wiadomości przeznaczonych dla właściciela kluczy. Klucz prywatny jest tajny i tylko przy jego pomocy można odszyfrować to, co zostało zakodowane kluczem publicznym. Najszerzej stosowanym algorytmem szyfrowania asymetrycznego jest **RSA** (*Rivest, Shamir, Adleman*), przedstawiony bardziej szczegółowo w [1].

System kryptograficzny z kluczem publicznym i prywatnym może być wykorzystywany do podpisywania dokumentów cyfrowych. Jednak w tym przypadku rola kluczy zostaje odwrócona. Ponieważ klucz prywatny przechowywany jest wyłącznie u podpisującego, służy on do szyfrowania danych. Klucz publiczny, ogólnie dostępny, służy do deszyfrowania i upewnienia się, czy tylko właściciel skorzystał z klucza prywatnego. Nadawca szyfruje dokument używając swojego klucza prywatnego. Odbiorca deszyfruje dokument używając klucza publicznego nadawcy weryfikując w ten sposób jego podpis. Podpis ten jest prawdziwy, gdyż został zweryfikowany przez użycie klucza publicznego nadawcy; podpis nie może być sfałszowany, gdyż tylko nadawca zna swój klucz prywatny. Podpisany dokument nie może być zmieniony, gdyż zmieniony dokument nie da się rozszyfrować kluczem publicznym nadawcy.

Wadą takiego sposobu podpisywania dokumentów jest to, że podpis jest co najmniej tak długi, jak sam dokument, co uniemożliwia praktyczne zastosowanie tej, jednak wymagającej dużych mocy obliczeniowych, procedury. Dlatego stosuje się procedurę z wykorzystaniem jednokierunkowej funkcji skrótu, tzw. funkcji *hash*.

Funkcja skrótu (haszująca)

W procedurze podpisu elektronicznego pierwszym krokiem jest wygenerowanie z pliku zawierającego dane, tzw. skrótu (*hash*). Jest to jednokierunkowe przekształcenie matematyczne zamieniające ciąg bitów dowolnej długości w inny ciąg bitów o zadanej długości (np. 128, czy 192 bity). Bezpieczna kryptograficznie funkcja skrótu powinna spełniać trzy podstawowe założenia:

– brak możliwości wygenerowania dwóch wiadomości o takim samym skrócie (z czysto matematycznego punktu widzenia jest to niewykonalne. Jeśli wyjście funkcji ma 128 bitów oznacza to, że biorąc $2^{128}+1$ różnych stanów wejściowych jakiś wynik na pewno się powtórzy),
– brak możliwości wygenerowania dwóch wiadomości o takim samym skrócie, czyli brak tzw. kolizji,
– brak możliwości odtworzenia danych wejściowych na podstawie skrótu.

Podkreślmy, że uznanie funkcji za bezpieczną do zastosowań kryptograficznych opiera się wyłącznie na *domniemaniu* odporności na znane ataki kryptoanalityczne, nie zaś na matematycznych dowodach gwarantujących niemożność złamania [2]. Istnienie jednokierunkowych funkcji nie zostało dotychczas dowiedzione. Poważne słabości znaleziono w wielu funkcjach skrótu, które historycznie uchodziły za bezpieczne, nawet w jeszcze używanych (m.in. w SHA, MD5).

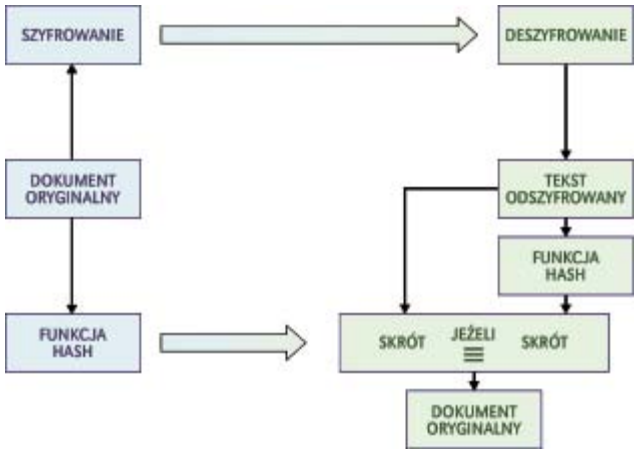
Funkcje haszujące, używane obecnie w kryptografii to MD5, SHA-1, SHA-2, RIPEMD-160. Jedną z najbardziej popularnych rodzin funkcji skrótu jest rodzina MD (*Message Digest*) Ronalda Rivesta (współtwórcy RSA). **MD5** (*Message-Digest Algorithm 5*), piąta wersja funkcji została opracowana w 1991 roku, która z dowolnego ciągu danych generuje 128-bitowy skrót. W 2004 r. znaleziono sposób na generowanie kolizji w MD5, co spowodowało, że nie jest już ona polecana do zastosowań wymagających wysokiego poziomu bezpieczeństwa [3]. Jest jednak w dalszym ciągu powszechnie stosowana w internecie, jako suma kontrolna przesyłanych plików.

SHA (*Secure Hash Algorithm*) to rodzina kryptograficznych funkcji skrótu zaprojektowanych przez NSA (*National Security Agency*) i publikowanych przez NIST (*National Institute of Standards and Technology*) [4]. Pierwsza z tych funkcji, opublikowana w 1993 r., została wycofana ze względu na oficjalnie nieujawnione wady. Została ona zastąpiona w 1995 r. przez algorytm **SHA-1**. Algorytm ten generuje 160-bitowy skrót z wiadomości o maksymalnym rozmiarze 2^{64} bitów, w budowie jest podobna do MD5.

Podobnie jak MD5, algorytm ten nie jest już zalecany do nowych aplikacji [5]. W 2001 r. powstały cztery następne warianty, określane jako **SHA-2** (SHA-224, SHA-256, SHA-384, SHA-512). Obecnie NIST prowadzi publiczny konkurs na następcę dotychczasowych funkcji skrótu. Na razie wiadomo tylko, że nowa funkcja zostanie nazwana **SHA-3** [6].

RIPEMD to funkcja skrótu opracowana w ramach projektu Unii Europejskiej o nazwie RIPE (*RACE Integrity Primitives Evaluation*) realizowanego w latach 1988-1992. W 1996 r. powstała wersja generująca skrót 160-bitowy nazwana **RIPEMD-160**. W 2004 r. Xiaoyun Wang, Dengguo Feng, Xuejia Lai oraz Hongbo Yu opublikowali dokument, w którym dwie pary wiadomości produkujących te same skróty [7]. Algorytm RIPEMD-160 jest stosunkowo mało popularny i słabo zbadany z punktu bezpieczeństwa stosowania.

Zastosowanie funkcji skrótu umożliwia więc następującą procedurę (*rys. 1*): nadawca najpierw tworzy skrót, a następnie podpisuje ten skrót szyfrując go kluczem prywatnym i przesyła wraz z dokumentem odbiorcy. Odbiorca używa tej samej funkcji hashującej do otrzymania skrótu dokumentu, a następnie deszyfruje podpisany skrót używając klucza publicznego nadawcy. Jeżeli zdeszyfrowany skrót zgadza się z przesłanym, to podpis jest prawdziwy. Zalety tej procedury są oczywiste: podpis jest znacznie krótszy od dokumentu, a jego wiarygodność można sprawdzić bez oglądania samego dokumentu. W praktyce sam przesyłany dokument jest również szyfrowany, ale już jakimkolwiek szybkim algorytmem symetrycznym, np. DES.



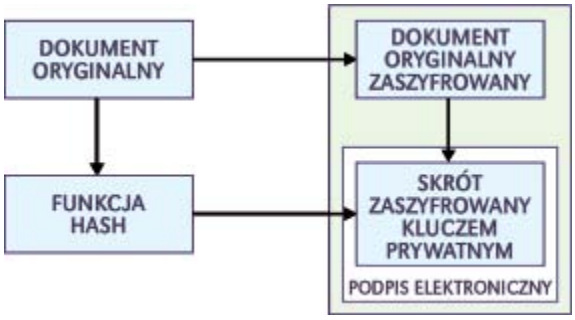
Rys. 1. Schemat weryfikacji integralności przesyłki przy użyciu funkcji skrótu wiadomości [8]
Fig. 1. Message integrity verification using a hash function [8]

Procedura złożenia i weryfikacji podpisu elektronicznego

Praktyczna procedura złożenia podpisu pod przygotowanym wcześniej dokumentem elektronicznym jest następująca: w pierwszym kroku obliczany jest skrót dokumentu, np. za pomocą funkcji SHA-1 i przesyłany do karty kryptograficznej. Tam wykonywane jest szyfrowanie tego skrótu, np. według algorytmu RSA za pomocą klucza prywatnego, zapisanego na tej karcie. Warunkiem wykonania tej operacji jest jej uwierzytelnienie kodem PIN. Wygenerowane dane odsyłane są do komputera i dołączane do oryginalnego dokumentu. Dodatkowo do dokumentu oraz zaszyfrowanego skrótu dołączany zostaje certyfikat zawierający dane osoby składającej podpis oraz jej klucz publiczny. Tak przygotowane dane można nazwać podpisem elektronicznym dołączonym do dokumentu i z nim powiązanym (*rys. 2*).

Weryfikacja dokumentu z podpisem polega na ponownym obliczeniu skrótu z dokumentu. W następnym kroku, przy pomocy klucza publicznego rozszyfrowywany jest skrót dołączony do dokumentu. Jeśli rozszyfrowany skrót jest równy obliczonemu skrótowi, wtedy weryfikacja jest pozytywna.

Zastosowanie technik kryptograficznych w procedurze podpisu elektronicznego nie wystarcza dla pewności, że przesłany dokument jest właściwy. Zaszyfrowanie lub elektroniczne podpisanie dokumentu nie daje bowiem gwarancji, że osoba która użyła klucza prywatnego jest tą, za którą się podaje. Gwarancję taką daje dopiero system certyfikacji kluczy.



Rys. 2. Konstrukcja przesyłki podpisanej elektronicznie przez nadawcę
Fig. 2. Document signed electronically by the sender

Certyfikat cyfrowy to elektroniczne zaświadczenie, że dane służące do weryfikacji podpisu elektronicznego są przyporządkowane określonej osobie i potwierdzają jej tożsamość. Certyfikacji dokonuje odpowiedni organ, poświadczający autentyczność danego klucza publicznego. Jest to tzw. Zaufana Trzecia Strona (*Trusted Third Party*). Zadaniem urzędu certyfikacji jest wydawanie i zarządzanie certyfikatami.

Certyfikat cyfrowy posiada następujące informacje: unikalny numer seryjny, tożsamość urzędu certyfikacji wydającego certyfikat, okres ważności certyfikatu, identyfikator właściciela certyfikatu (imię, nazwisko, pseudonim, e-mail itp.), klucz publiczny właściciela certyfikatu i podpis cyfrowy urzędu certyfikacji potwierdzający autentyczność certyfikatu.

Dodatkowymi aspektami, o których warto w tym momencie wspomnieć są: ważność certyfikatu oraz możliwość jego zawieszenia lub unieważnienia. W ustawie o podpisie elektronicznym określona jest maksymalna ważność certyfikatu kwalifikowanego, która wynosi 2 lata. Po tym czasie certyfikat staje się nieważny i podpisy złożone po upłynięciu terminu ważności stają się automatycznie weryfikowane negatywnie. Możliwe jest także unieważnienie lub zawieszenie certyfikatu, następuje ono na wyraźną prośbę użytkownika np. w przypadku poznania przez osoby niepowołane kodu PIN. Fakt ten zostaje odnotowany i opublikowany na tzw. liście CRL (*Certificate Revocation List*). Listę taką publikuje i uaktualnia codziennie na swoich stronach każdy kwalifikowany urząd certyfikacji.

Na polskim rynku działa sześć firm oferujących zestawy do składania bezpiecznego podpisu elektronicznego. Są to Krajowa Izba Rozliczeniowa, Unizeto, Polska Wytwórnia Papierów Wartościowych, Mobicert, Enigma oraz Safe Technologies (CenCert). W skład typowego zestawu wchodzą: czytnik kart kryptograficznych, karta kryptograficzna oraz zapisany na karcie certyfikat, który zawiera parę kluczy RSA, a także informacje o osobie na którą jest wystawiony. Ponieważ bezpieczny podpis elektroniczny weryfikowany jest tzw. certyfikatem kwalifikowanym, ma moc prawną odpowiadającą podpisowi odręcznemu i przy jego wystawianiu i wydawaniu weryfikowana jest tożsamość osoby. Niezbędne jest także odpowiednie oprogramowanie służące do składania oraz weryfikacji podpisów elektronicznych.

Warto w tym momencie zaznaczyć, że taka konstrukcja systemu powoduje, iż dane służące do składania bezpiecznego podpisu faktycznie znajdują się pod wyłączną kontrolą podpisującego. Certyfikat z parą kluczy znajdują się na karcie kryptograficznej i dostęp do nich chroniony jest kodem PIN. Niemożliwe jest także skopiowanie lub usunięcie kluczy z karty. Po kilkukrotnym błędnym wpisaniu kodów PIN oraz PUK dostęp do certyfikatu oraz kluczy zostaje zablokowany.

Bezpieczeństwo podpisu elektronicznego

Algorytm RSA jest dobrze znany, a jego wady są bardzo dokładnie opisane. Oparty jest na złożoności faktoryzacji dużych liczb, dlatego jedynym zabezpieczeniem przed złamaniem klucza prywatnego jest czasochłonność obliczeniowa. Na dziś największym kluczem, jaki praktycznie udało się rozłożyć na czynniki pierwsze jest klucz 768-bitowy. W Polsce stosowane są klucze 1024-bitowe, co, jakkolwiek wydaje się niewiele więcej, to zapewnia wystarczający margines bezpieczeństwa. Prawdopodobnie następną minimalną stosowaną długością klucza będzie 2048 bitów.

Aktualnie stosowana funkcja SHA-1, także jest uznawana za bezpieczną w istniejących aplikacjach. Znane są co prawda próby ataku, jednak na razie są zbyt mało obliczeniowo skuteczne, aby można było mówić o realnym zagrożeniu dla bezpieczeństwa podpisu elektronicznego. Ponadto coraz bardziej powszechnie stosowana rodzina funkcji SHA-2, stwarza możliwość zastąpienia SHA-1 bezpieczniejszym algorytmem.

Użycie kombinacji algorymtów RSA oraz SHA-1, zabezpiecza więc skutecznie przed ingerencją w podpisany dokument. Każda jego modyfikacja powoduje zmianę skrótu, co doprowadza do negatywnej weryfikacji podpisu. Żeby skutecznie podrobić podpis należałoby wygenerować drugi identyczny skrót, co jest na dzień dzisiejszy praktycznie niewykonalne. Poza tym, nie wystarczy przełamanie tylko jednego zabezpieczenia. Istotne jest na przykład, aby drugi „podrobiony" dokument niewiele różnił się od oryginału. W końcu, jeśli nawet będą istnieć dwa dokumenty o takim samym skrócie, muszą być te skróty podpisane, a do tego niezbędna jest znajomość klucza prywatnego, którego odtworzenie wymaga ogromnej złożoności obliczeniowej. Poznanie tego klucza z karty kryptograficznej również nie jest proste, gdyż karty na których przechowywane są klucze oraz certyfikaty kwalifikowane są specjalnie zabezpieczane przed dostępem do pamięci osób trzecich.

Podsumowanie

Podpis elektroniczny stanowi ukoronowanie osiągnięć kryptografii naszych czasów wykorzystującej zarówno szyfrowanie asymetryczne, jak i jednokierunkową funkcję skrótu. Jednak każdy system kryptograficzny jest tak słaby, jak jego najsłabsze ogniwo. Tym najsłabszym ogniwem jest człowiek. Często każdy użytkownik różnych kart bankowych, kredytowych i in. musi pamiętać wiele kodów PIN. Niektóre osoby przechowują więc numery PIN zapisane na kartce przylepionej do monitora lub leżącej na biurku. Równie często zestawy do składania podpisu są na stale podłączone do komputera użytkownika i łatwo dostępne. Taka kombinacja niedopatrzeń pozwala składać podpisy osobie postronnej. Każdy, kto używa podpisu elektronicznego powinien o tym pamiętać.

Literatura

[1] Nowakowski W.: Algorytm RSA – podstawa podpisu elektronicznego. Elektronika 6/2010, Warszawa 2010.
[2] Sklarov D. V.: Łamanie zabezpieczeń programów. RM, Warszawa 2004.
[3] http://blog.securitystandard.pl/news/342130.html
[4] http://csrc.nist.gov/groups/ST/toolkit/index.html
[5] http://tools.ietf.org/html/rfc3174
[6] http://ipsec.pl/kryptografia/2008/trzydziestu-kandydatow-do-sha-3.html
[7] http://eprint.iacr.org/2004/199.pdf
[8] http://krystian.jedrzejczak.webpark.pl/podpis.htm#podpis
[9] Tanaś R.: http://zon8.physd.amu.edu.pl/~tanas/
[10] http://www.ebanki.pl/technika/podpis_cyfrowy.html
[11] RSA Laboratories. RSA Laboratories' Frequently Asked Questions About Today's Cryptography, Version 4.1. RSA Security Inc. 2002.
[12] Ustawa z dnia 22 sierpnia 2001 roku o podpisie elektronicznym (Dz.U. 2001 nr 130 poz. 1450, tekst ujednolicony).