Eliptické křivky a šifrování (2.)

V předchozím dílu jsme se seznámili s eliptickými křivkami, nyní se posíváme na jejich využití k elektronickému podpisu a k šifrování a řekneme si pár poznámek k jejich bezpečnosti. Upozorníme také na různé standardy, v nichž se do detailů dozvíte, jak tyto systémy využít.

Eliptické křivky nad tělesem GF(2m)

Zatím jsme poznali eliptickou křivku nad tělesem GF(p). U tělesa GF(2m) je situace složitější jen pro matematiky a programátory, jinak je podstata stejná jako u GF(p). Protože bychom zde všechny rozdíly a jejich důvody stejně nemohli rozebrat (jiná rovnice křivky, jiné souřadnice, přibývá reprezentace prvků tělesa v různých bázích, jinak se definuje součet bodů na křivce atd.), spokojíme se pro další výklad s GF(p).

Šifrování s ECC

V čem je tedy podstata šifrování pomocí ECC (Elliptic Curve Cryptosystem)? Ukážeme si ji na analogii Diffie-Hellmanova schématu výměny klíče. Tento algoritmus řeší situaci, kdy si dvě strany, A a B, chtějí vyměnit tajnou informaci přes veřejný kanál. Jak je to u všech systémů s veřejným klíčem nutné, i zde se předpokládá, že každá ze stran má k dispozici důvěryhodnou cestou získaný veřejný klíč protistrany. Navíc zde předpokládáme, že obě strany sdílejí stejnou křivku E a stejný bod P Ž E. Označíme-li po řadě dA a QA privátní a veřejný klíč strany A, a obdobně dB a QB pro stranu B, potom obě dvě strany si mohou ustavit společný klíč - bod Z na křivce E, aniž spolu komunikují. Strana A vypočte bod Z jako dAQB a strana B jako dBQA Tyto body jsou skutečně stejné, neboť Z = dAQB = dA(dBP) = (dAdB)P a současně Z = dBQA = dB(dAP) = (dBdA)P. Tedy každá strana vezme veřejný bod (klíč) protistrany a sečte ho n-krát, kde n je její privátní klíč. Protože obě strany vycházejí ze stejného bodu P, dospějí zákonitě do stejného bodu Z. Tento bod ovšem nezná nikdo jiný než ony, v čemž je podstata ustavení společného tajného prvku. Jak ho použijí k utajení komunikace, je jiná otázka. Obvykle se z x-ové souřadnice bodu Z odvozují klíče na sezení pomocí různých dalších technik a pomocí klíče na sezení a vybraného symetrického algoritmu se příslušné spojení šifruje. Ani v případě, že by se na komunikačním kanálu předávaly i hodnoty QA a QB, není bod Z prozrazen, protože útočník z nich není schopen určit privátní hodnoty dA a dB díky složitosti diskrétního logaritmu (viz minulý díl). Přenos těchto hodnot přichází v úvahu například tehdy, když jedna z komunikujících stran nemá veřejný klíč založený na stejné křivce a bodu jako protistrana. Potom si odesílatel (například e-mailu) vezme bod a křivku protistrany a "ad hoc" si vytvoří svůj klíčový pár s touto křivkou. Svůj bod Q pak společně se zašifrovanou zprávou pošle protistraně. Ani v tomto případě tedy obě strany nemusí být spojeny on-line.

Elektronický podpis s ECC

Uveďme si nyní, jak definuje elektronický podpis pomocí eliptických křivek (ECDSA - elliptic curve digital signature algorithm) standard FIPS 186-2, který zmiňuje i naše vyhláška k zákonu o elektronickém podpisu. Standard definuje více křivek, zde si vybereme tu nad tělesem GF(p) s nejmenším prvočíslem p (192bitovým). Parametry křivky vidíte v rámečku "Křivka P-192", v dalším samostatném rámečku jsou uvedeny příslušné postupy. Kdo chce hlouběji porozumět důvodům takovéto definice, měl by se podívat na definici podpisového schématu DSA (viz literatura). Toto schéma (multiplikativní grupa) se pak transformuje na eliptickou křivku (aditivní grupa) tak, že operace násobení prvků g * g * g * g *.... (tj. gk) se převede na sčítání bodů na křivce P + P + P + P + ... (tj. kP).

Standardy a literatura

Často citovanou normou pro digitální podpis je FIPS 186-2 (http://csrc.nist.gov/fips/), která zrovnoprávňuje podpis na bázi RSA (viz lit.), DSA (viz lit.) i ECDSA (eliptická varianta DSA). ECDSA vychází z normy ANSI X9.62. Ta, stejně jako FIPS 186-2, pak těží z práce skupiny P1363 organizace IEEE, která definuje řadu asymetrických algoritmů, včetně těch na bázi eliptických křivek (http://grouper.ieee.org/1363/ index.html). ECC se zabývá i ANSI norma X9.63. Další skupinu tvoří různé normy ISO používající ECC: například ISO 14888-3 definuje digitální podpis, ISO/IEC 15946 definuje podpisy, šifrování a výměnu klíče, ISO/IEC 9798-3 autentizaci a ISO/IEC 11770-3 klíčové hospodářství. Dále jsou k dispozici různé internetové standardy IETF, využívající eliptické křivky pro internetové použití

(http://www.ietf.org/), standardy WAP fóra pro bezdrátové komunikace, zejména mobilní telefony (například Wireless Transport Layer Security, http://www.wapforum.org). ECC prosazuje také komerční uskupení SECG, vydávající standardy na bázi ECC (http://www.secg.org), a to jak pro digitální podpisy, tak na jejich využití k elektronickému podpisu a k šifrování a řekneme si pár poznámek k jejich bezpečnosti. Upozorníme také na různé standardy, v nichž se do detailů dozvíte, jak tyto systémy využít. pro šifrování. ECC zaujme zejména u mobilních telefonů, kde vystupuje do popředí příznivý poměr cena/výkon, dlužno ale poznamenat, že v rámci honby za výkonem se zde definuje ECC s nízkou až velmi nízkou bezpečností. Velmi dobrým začátkem pro studium jak vlastních eliptických křivek, tak norem je web společnosti Certicom, kde pracují uznávaní kryptologové (http://www.certicom.com).

Bezpečnost

V tabulce "Doporučené délky klíče podle NIST" vidíte také porovnání bezpečnosti symetrických systémů, u nichž se předpokládá útok hrubou silou (vyzkoušení všech možných klíčů), a bezpečnosti eliptických křivek, kde se uvažuje složitost řešení problému diskrétního logaritmu pomocí Pollardovy ró metody. Tuto tabulku zpracoval NIST jako doporučení pro federální použití v USA. V návrhu nově připravovaného dokumentu "Příručka klíčového hospodářstvíŽ z 3. 7. 2002 pak NIST zpřesňuje délky u ECC tak, že uvádí logičtější požadavek na "velikost řádu generujícího boduŽ (http://csrc.nist.gov/encryption/kms/guideline-1.pdf, tabulky 8 a 9). Třetí a čtvrtý sloupec v naší tabulce berte proto spíše orientačně.

Shrnutí

Kryptografie eliptických křivek (ECC) je nový nadějný obor. Eliptické křivky poskytují dobrý poměr cena/výkon a stávají se součástí nejdůležitějších světových standardů (ANSI, ISO, IETF). Implementaci těchto nástrojů nic nebrání, snad jen jejich nezvyklost. Také jejich bezpečnosti se věnuje značná pozornost, takže obavy tohoto druhu asi nebudou tím hlavním důvodem, proč eliptické křivky nejsou masově používány a "staré dobréŽ algoritmy RSA, DH a DSA ještě nevyklízejí pole. Každopádně však tam, kde jsou k dispozici jen omezené hardwarové zdroje, nemají eliptické křivky konkurenci.

Vlastimil Klíma, autor@chip.cz

Literatura

[1] Klíma, V.: DSA: Podpis bez pera i papíru, Chip 5/99, str. 40 - 42
[2] Klíma, V.: Bude nás podepisovat RSA?, Chip 9/00, str. 50 - 52
[3] Klíma, V.: SHA-1: Výživná haše, , Chip 3/99, str. 40 - 43
[4] Elektronický archiv uvedených i dalších článků:
http://www.decros.cz/bezpecnost/_kryptografie.html

Křivka P-192

E : y2 ľ x3 - 3x + b (mod p) prvočíselný modul p = 6277101735386680763835789423207666416083908700390328520][28243][45111][8451][25956][17221][7415][45824][4913][7683][43023][7939][63733][45824][11264] prvočíselný řád křivky #E = 6277101735386680763835789423176059013767194773182828520][28243][45111][8451][25956][17221][7415][45824][4913][7683][43023][7939][63733][45824][11264] kofaktor = 1, protože řád křivky je prvočíslo; b = 64210519 e59c80e7 0fa7e9ab 72243049 feb8deec c146b9b1 (hex.); (generující) bod P: xP = 188da80e b03090f6 7cbf20eb 43a18800 f4ff0afd 82ff1012 (hex.), yP = 07192b95 ffc8da78 631011ed 6b24cdd5 73f977a1 1e794811 (hex.).

Generování klíče pro ECDSA

Vybereme eliptickou křivku E nad GF(p). Počet bodů křivky #E by měl být dělitelný velkým prvočíslem n. Zvolíme bod P Ž E řádu n (poznamenejme, že ANSI X9.62 požaduje, aby řád křivky byl větší než 2160). Vybereme jedinečnou a nepredikovatelnou hodnotu privátního klíče, číslo d Ž [1, n-1]. Vypočteme veřejný bod Q = dP. Veřejný klíč tvoří čtveřice (E, P, n, Q).

Vytvoření podpisu pomocí schématu ECDSA

Mějme zprávu m.
Vybereme jedinečné a nepredikovatelné číslo k Ž [1, n - 1].
Vypočteme bod kP = (x1, y1) a číslo r = x1 mod n.

Je-li r = 0, pak postup opakujeme od generování čísla k (to je nutné proto, aby v hodnotě s byl obsažen privátní klíč, viz dále).

Vypočteme k-1 mod n.

Vypočteme s = k-1 {h(m) + dr} mod n, kde h je hašovací funkce SHA-1 (viz literatura).

Je-li s = 0, pak opět jdeme na první bod - generování nového k (neexistovalo by s-1 mod n, viz dále proces ověření).

Podpisem zprávy m je dvojice čísel (r, s).

Ověření podpisu ECDSA

Mějme zprávu m a její podpis (r, s).

Důvěryhodným způsobem získáme veřejný klíč podepisujícího (E, P, n, Q).

Ověříme, že r, s jsou z intervalu [1, n-1].

Vypočteme w = s-1 mod n a h(m).

Vypočteme u1 = h(m)w mod n a u2 = rw mod n.

Vypočteme u1P + u2Q = (x0, y0) a v = x0 mod n.

Podpis je platný právě tehdy, když v = r.