Kryptologie pro praxi – DSA, ECDSA

Nejznámějšími schématy pro digitální podpisy jsou v současné době RSA, DSA a ECDSA. Podpisové schéma DSA (Digital Signature Algorithm) definované standardem [3] a jeho varianta Elliptic Curve DSA (ECDSA), vzniklá přenesením DSA nad algebru bodů rovinné eliptické křivky [3] a [4], hrají právem důležitou úlohu v současné asymetrické kryptografii. Na území USA jsou totiž DSA a ECDSA schváleny standardem [3] pro použití ve vládních institucích, což je určitá výsada, kterou donedávna nemělo ani RSA [2]. Teprve novelizace zmíněného standardu v roce 2000 v tomto smyslu RSA a DSA zrovnoprávnila. DSA je ovšem zajímavé i z didaktického hlediska a to zejména pro oblast aplikované kryptografie, ve které se tomto seriálu pohybujeme. Lze na něm totiž dobře ilustrovat, že i zjednodušování kryptografie vyžadované aplikačními inženýry má své meze. Algoritmus DSA totiž jednak slouží pouze pro podepisování (čímž boří častý mýtus, že asymetrická šifra a podpis jsou jedno a totéž), jednak je vnitřně konstruován zcela odlišně od RSA. V DSA proto nenajdeme žádné mechanizmy typu šifrovací/odšifrovací transformace a princip ověření pravosti podpisu je od základu jiný než u podpisového schématu vzniklého použitím reverzibilní asymetrické šifry (viz [1]). I přesto se bohužel dodnes vyskytují texty, které se snaží DSA coby „převlečenou šifru“ popisovat.

Klíče, transformace a podpis

Veřejnými parametry DSA je trojice celých čísel (p, q, g), kde p je 1024bitové prvočíslo určující multiplikativní grupu Z_p^*, q je 160bitový prvočíselný dělitel čísla $p-1$ a g je číslo z intervalu $\langle 2,p-2\rangle$, které má v Z_p^* řád q, tedy $g^q \bmod p=1$. Algoritmus pro generování veřejných parametrů je podrobně popsán v [3]. Privátní klíč x je libovolné celé číslo z intervalu $\langle 1,q-1\rangle$. Jemu odpovídající veřejný klíč je vypočítán jako $y=g^x \bmod p$. Veřejné parametry (p, q, g) a klíč y jsou určeny ke zveřejnění (formou certifikátu, atp.), privátní klíč x je držen v tajnosti. Schéma umožňuje také sdílet konkrétní sadu veřejných parametrů, například mezi více uživateli téhož systému.

Podpis zprávy M probíhá následovně: Nejprve se vypočte její otisk $h=SHA\text{–}1(M)$ (hašovací funkce SHA–1 viz *ST* 2/2004). Vlastní podpis je reprezentován dvojicí celých čísel (r, s) vypočtených jako: $r=(g^k \bmod p) \bmod q$, $s=(h+xr)k^{-1} \bmod q$, kde $kk^{-1} \bmod q=1$ a k je náhodně vygenerované číslo z intervalu $\langle 1,q-1\rangle$, které musí být z bezpečnostních důvodů generováno vždy znovu pro každý podpis a musí být dokonale utajeno (nejlépe je ho po podpisu ihned zničit). Špatné generování a zacházení s k může vést k závažnému oslabení celého schématu. Pokud by například v extrémním případě došlo k prozrazení celé hodnoty k, lze s pouhou znalostí veřejných parametrů a podepsané zprávy triviálně přímo z hodnoty podpisu vypočítat privátní klíč. Proto také bývá k nazýváno jako klíč zprávy. Poznamenejme, že čísla r a s jsou obě z intervalu $\langle 1,q-1\rangle$, takže podpis DSA lze vždy formátovat jako 320bitový řetězec (2×160 b).

Ověření podpisu (r,s) zprávy M probíhá takto (vynecháváme zde některé rutinní integritní kontroly, viz [3]): Ověřující strana nejprve opět určí otisk $h=SHA\text{–}1(M)$. S jeho využitím vypočte $v=(g^u y^w \bmod p) \bmod q$, kde $u=(hs^{-1}) \bmod q$, $w=(rs^{-1}) \bmod q$ a $(ss^{-1}) \bmod q=1$. Podpis je platný právě tehdy, když platí $v=r$. Vidíme, že v celém procesu skutečně nedošlo k žádnému šifrování (a už vůbec ne odšifrování) k získání zprávy M ani jejího otisku, neboť hodnota v (rovnající se r, pokud vše dopadlo dobře) nemá s výsledkem takové operace nic společného (například r na M vůbec nezávisí).

Popis instance a transformací ve schématu ECDSA, které zde s ohledem na rozsah článku vynecháváme, je uveden v [4]. Co do principu, je zde situace přesně analogická DSA. V důsledku odlišné algebry se ovšem význam parametrů instance a zápis jednotlivých operací na první pohled znatelně liší. Při hlubším prostudování ovšem zmíněná analogie plně vynikne, v důsledku čehož se schéma stává podstatně srozumitelnějším.

Bezpečnost (EC)DSA

Bezpečnost DSA se matematicky opírá o takzvaný problém diskrétního logaritmu [6]. Jedná se o úlohu, kterou lze zadat k řešení nad mnoha algebraickými strukturami. S měnící se strukturou se přitom kromě konkrétního popisu úlohy mění i možné způsoby řešení a jejich složitost. V okamžiku vzniku DSA se jako optimální struktura jevila multiplikativní grupa celých čísel modulo prvočíslo p, respektive její jistá podgrupa prvočíselného řádu q, generovaná prvkem g (viz výše). I dnes a v blízké budoucnosti můžeme tuto strukturu s ohledem na bezpečnost a praktickou zvládnutelnost operací považovat za vyhovující. Kdyby se však přece jen kolem této struktury začaly časem objevovat nějaké pochybnosti, existuje zde dobře zpracovaná alternativa v přechodu na aditivní grupu bodů rovinné eliptické křivky [4], [6]. Za cenu vyšších implementačních nároků tak získáme schéma, jehož prolomení se při odpovídající délce klíče zdá být významně obtížnější. Z pohledu praxe se tento krok chápe takto: Buďto jej využijeme pro zvýšení bezpečnosti (a délku parametrů a klíčů zhruba zachováme), nebo jej použijeme pro snížení délky kritických parametrů (při zachování úrovně bezpečnosti). Dodejme, že zatím problematika ECDSA plní spíše přednášky konferencí a na svou významnější praktickou roli srovnatelnou s úlohou RSA a DSA teprve čeká.

Stejně jako v případě RSA, představují pro (EC)DSA v současnosti největší riziko chyby vzniklé při jeho nepozorné implementaci. Jako ilustrační příklad může sloužit útok [5]. Zde konkrétně byl mj. podceněn význam integrity veřejných parametrů DSA. Architekti domnívajíce se snad, že tady není co chránit, je totiž nechali „bezprizorně“ ležet na disku počítače hned vedle pečlivě zašifrované hodnoty privátního klíče. Tím ovšem umožnili útočníkovi jejich snadnou výměnu, o které bylo ukázáno v [5], že ji lze triviálním způsobem využít k útoku. Stačí jen dosadit vhodné hodnoty a pak počkat, až nic netušící uživatel vypustí do světa první nepovedený podpis, ze kterého se už snadno vypočítá tolik střežený privátní klíč. Poučením z této konkrétní kauzy je, že u veřejných parametrů (a konečně i u privátního klíče) je nutné zajistit kvalitní kontrolu integrity. Zobecněným poučením pak je, že u asymetrických schémat obzvlášť záleží na každém detailu jejich implementace a ani chvilková nepozornost se zde nevyplácí.

Vlastimil Klíma, Tomáš Rosa,
klima@lec.cz, trosa@ebanka.cz

LITERATURA

[1] Kryptologie pro praxi – asymetrické metody, ST č. 8/2003

[2] Kryptologie pro praxi – metoda RSA, ST č. 3/2004

[3] FIPS PUB 186-2: Digital Signature Standard (DSS), NIST USA, 2000-2001

[4] Johnson, D., Menezes, A., Vanstone, S.: The Elliptic Curve Digital Signature Algorithm (ECDSA), International Journal of Information Security, Vol 1, Issue 1, pp. 36-63, Springer-Verlag, 2001

[5] Klíma, V., Rosa, T.: Attack on Private Signature Keys of the OpenPGP Format…, IACR ePrint report 2002/076, http://eprint.iacr.org/2002/076/

[6] Archivy článků: http://cryptography.hyperlink.cz a http://crypto.hyperlink.cz