RSA afkóðun

Smári P. McCarthy

6. nóvember 2005

Efnisyfirlit

1 Ágrip af sögu RSA

Whitfield Diffie, Martin Hellman og Ralph Merkle birtu Diffie-Hellman-Merkle dulmálskerfið árið 1976. Dulmálskerfið þeirra var þeim eiginleikum gætt að í stað eins dulmálslykils voru þeir tveir, og að dulmálskerfið var ósamhverft. Diffie og Hellman voru að íhuga einkalykladulmálsfræði (*Public Key Cryptography*), en Merkle var aðallega að skoða aðferðir til þess að dreifa dulmálslyklum. Þegar að þeir áttuðu sig á því að það væri margt skylt með vinnu hópanna þá samtvinnuðust hugmyndir þeirra þriggja, og þær birtust í grein að nafni *New Directions in Cryptography* í tímaritinu IEEE Transactions on Information Theory, sem var þó bara eftir Diffie og Hellman, og hefur kerfið verið kennt við þá síðan.

Hugmyndin þeirra var að það væri hægt að búa til stærðfræðilegar "fallgildrur", þ.e. aðferðir sem tættu gögn upp án þess að glata þeim, þannig að með réttu aukaílagi væri auðvelt að andhverfa þær, en annars væri það mjög torleyst.

RSA dulkóðunarreikniritið var afrakstur verka þriggja manna, Ronald Rivest, Adi Shamir and Leonard Adleman. Þeir birtu sameiginlega ritgerðina *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems* árið 1977, en í henni lögðu þeir fram RSA dulmálið. Þeir byrjuðu að vinna að RSA eftir að Diffie og Hellman birtu sitt dulmálskerfi.

Rivest, Shamir og Adleman töldu að þó svo að aðferð Diffie-Hellman væri góð, þá væri hún ekki nógu góð, þar sem að það kerfi bauð eingöngu upp á dulkóðun, en ekki stafrænar undirskriftir. Þeir skoðuðu um fjörtíu mismunandi dulmálsaðferðir og römbuðu svo niður á notkun á margfeldi prímtalna sem einátta fall (*trapdoor one-way function*).

RSA dulmálið var fyrst auglýst í ágúst 1977, í dálki í Scientific American þar sem að kerfinu var lýst í grófum dráttum. Þar buðust þeir félagar einnig til þess að senda hverjum þeim sem hafði áhuga afrit af ritgerð sinni, sem fór nokkuð fyrir brjóstið á þjóðaröryggisstofnun Bandaríkjanna, NSA, sem sá hversu öflugt þetta kerfi var og hversu hættulegt það yrði ef að þetta félli í óvinahendur. Þeir kröfðust þess að hætt yrði dreifingu dulmálskerfisins um leið, en það rann á daginn að ekki voru lagalegar stoðir til staðar fyrir slíkri kröfu. Þegar að greinin var svo birt í IEEE Transactions on Information Theory var leyndarmálið komið í mjög almenna umferð, og allar frekari tilraunir NSA til þess að hindra dreifingu þess voru úr sögunni.

Árið 2000 rann út einkaleyfi þeirra félaga á RSA dulmálskerfinu í Bandaríkjum Ameríku, með þeim afleiðingum að nú geta allir heimsins íbúar notast við dulmálskerfið. Margt hefur breyst síðan að RSA var birt, enda er búið að uppgötva marga galla og margar holur í reikniritinu sem gera það

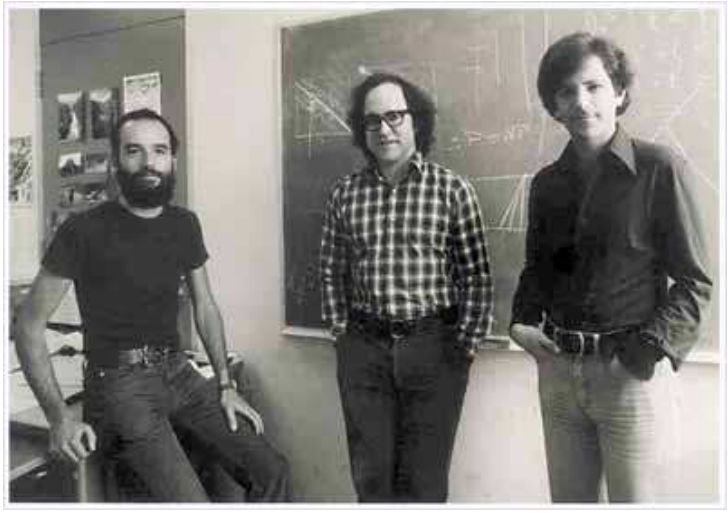
óöruggt. Ennfremur er búið að renna traustari stoðum undir einkalykladulmálskerfi með fyrirbærum á borð við OAEP, AES og Rijendel, ásamt því að RSA er komið í mun meiri almenna notkun með tilkomu PGP (*Pretty Good Privacy*) kerfisins eftir Philip Zimmermann.

Ljóst er að RSA er ekki óbrjótanlegt dulmál, heldur er það frekar torbrjótanlegt - þ.e., að á meðan að aðferðir manna til þess að þátta tölur í frumþætti hafa ekki einfaldari tímaflækju en þær hafa í dag, þá er alltaf hægt að auka þannig á stærð lyklanna að talan n sé óþáttanleg innan raunhæfs tímaspans. Fræðimenn hafa nefnt tvo mjög góða möguleika til þess að gera RSA auðleysanlegt:

1. Finna nýja aðferð til þess að leysa þáttunarvandamál með $O(n)$ eða, ákjósanlega, $O(\log n)$ tímaflækju.

2. Með skammtatölvu ætti, samkvæmt kenningunni, að vera hægt að framkvæma allar hugsanlegar reikniaðgerðir á augnabliki. Það hefði í för með sér að þáttunarvandamál, hversu mikil sem stærðargráðan á n er, yrðu nær óendanlega fljótleyst.

Vegna ótta manna við að (2) yrði að raunveruleika, þá var þróuð skammtafræðilega örugg aðferð til þess að skiptast á lyklum, og yrði þá hægt að notast við OTP (*One Time Pad*) dulkóðun í stórum stíl, sem yrði fullkomnlega öruggt, en útlistun á þeim aðferðum er utan sviðs þessarar ritgerðar.



Mynd 1: Rivest, Shamir og Adleman

2 RSA reikniritið

Til þess að notast við RSA reikniritið þarf að búa til lykilapar, þ.e. almenningslykil (*public key*) og einkalykil (*private key*). Þetta er gert í fimm skrefum:

1. Velja tvær stórar prímtölur p og q af handahófi, þannig að $p \neq q$.
2. Reikna $n = pq$.
3. Reikna $\phi(n) = (p-1)(q-1)$, þar sem ϕ er Eulers ϕ fallið [1]
4. Velja heiltölu e, þannig að $1 < e < \phi(n)$, sem er ósamþátta $\phi(n)$.
5. Reikna d þannig að $de \equiv 1 \pmod{\phi(n)}$.

Þá er einkalykillinn parið (n, d) og almenningslykillinn er parið (n, e). Þá sést að d er leyndarmálið í þessu, enda er ekki hægt að finna út d nema með því að reikna p og q út frá n, og finna svo rétt k þannig að $d = \frac{k\phi(n)+1}{e}$.

Reikniritið er svo skilgreint þannig:

$$I^l = O \pmod{n} \tag{2.1}$$

þar sem að I er inntak, O er úttak og l, lykillinn, er d ef að inntakið er dulkóðað en e ef dulkóða á inntakið. Dulkóðun fer þá fram með $T^e = C \pmod{n}$, þar sem að T er textinn sem dulkóða á og C er dulkóðað úttak, og afkóðun fer fram með

$$\begin{aligned} C^d &= (T^e)^d \\ &= T^{ed} \\ &= T^{k\phi(n)+1} \\ &= (T^{\phi(n)})^k T \\ &= 1^k \cdot T \pmod{n} \\ &= T \pmod{n}. \end{aligned}$$

Sem byggist á Eulersreglu $T^{\phi(n)} = 1 \pmod{n}$, sem gengur bara upp ef að T er ósamþátta n, sem er óhugnanlega líklegt.

[1] Eulers $\phi(n)$ fallið skilar fjölda talna minni en n sem eru ósamþátta við n.

3 Aðferðir við krökkun

Margar misgóðar aðferðir eru til að brjóta RSA. Flestar fela í sér *brute force* frumþáttun á n, en ekki eru allir vegir jafn færir í þessum efnum. Hér eru nokkrar aðferðir tíundaðar sem hafa nýst vel í gegnum tíðina.

3.1 Blokkaleit

Einn hugsanlegra galla við einkalykilsdulkóðun er að allir þurfa að hafa aðgang að reikniritinu sem notaður er til þess að dulkóða gögnin. Séu skilaboðin smá er með auðveldum hætti hægt að prófa að dulkóða öll möguleg skilaboð í einni blokk með almenningslykilnum uns samsvörun er fundin við blokk af dulkóðuðum texta. Þessi aðferð er þó sjaldnast hagnýt, þar sem að skilaboðin eru yfirleitt stærri en svo að þessi aðferð virki vel. Hagnýt útfærsla á þessu (*Adaptive chosen ciphertext attack*, Daniel Bleichenbacher 1998) var útfærð fyrir PKCS#1 (*Public Key Cryptography Standard #1*) dulkóðunarskemað, sem felur í sér sértækt blokkaskipan. Aðferðin gengur út á að reyna að finna fyrirsjáanlegar breytingar á dulkóðuðum texta út frá mismunandi frumtexta, og virkar eingöngu þegar að kóðunarskemað gefur af sér fyrirsjáanlegar blokkir. Auðvelt er að koma í veg fyrir slíkar árásir með OAEP (*Optimal Asymmetric Encryption Padding*), bilunarskema sem er byggð á Fiestel netum.

3.2 Að giska á d

Önnur aðferð er að nota þekktan streng til þess að giska á gildið á d. Þá er strengur, t.d. *HALLO*, valinn og dulkóðaður, og svo er notast við það að bæði dulkóðaða og ódulkóðaða afbrigði textans eru þekkt til þess að reyna að giska á d.

Þá væri hægt að giska á öll hugsanleg gildi á d uns dulkóðaði textinn afkóðast rétt, Það er því miður óttalega seinvirkt, og ekki gott ef að vitað er að n er mjög stór tala. Eftir að d hefur verið fundið er mjög einfallt að finna n út frá d, þar sem $de \equiv 1 \pmod{\phi(n)}$. Sýnt var fram á það árið 1990 að sé p á bilinu $]q...2q[$ og $d < \frac{n^{1/4}}{3}$ er hægt að finna d með fljótvirkum hætti út frá n og e.

3.3 Aðrar aðferðir

Nokkrar aðrar aðferðir fela í sér nýtingu á göllum, t.d. þegar að e og d eru mjög lágar tölur, þegar e er sameiginleg með mörgum lyklum innan fyrirtækis (sem gerir það auðveldara að giska á d), eða þegar að n er of lág. Það að n

er eingöngu 257 bitar að lengd í tilvikinu sem mér var falið að leysa gerir gæfumuninn, þar sem að 257 bita RSA lykill er auðþáttanlegur á nútíma tölvum.

4 Frumþáttun á n og ákvörðun d

Aðferðin sem ég ákvað að nota var *brute force* aðferðin, þ.e. að frumþátta n, sem er hluti af dreifilyklinum (n, e). Frumþáttunaraðferðir eru fjölmargar og mishraðar. Nokkrar helstu eru reiknirit Dixon's, Continued fraction factorization (CFRAC), Quadratic sieve, General number field sieve og Lenstra elliptic curve factorization. Síðastnefnda aðferðin notast við sporbauga sporaskja til þess að finna frumþætti, og er gríðarlega hraðvirk. Forritið GP PARI notast við þá aðferð og fjölmargar aðrar til þess að skila niðurstöðu á tiltölulega fljótlegan hátt, en forritið velur aðferð eftir eðli tölunnar. Frumþáttun á n skilaði sér eftir um klukkustundar keyrslu með Lenstra reikniritinu. Ég keyrði þáttunina samhliða á nokkrum tölvum með misjöfnum aðferðum til þess að reyna að tryggja að niðurstaða kæmi fyrir vikulok. Það að þáttunin var komin í hús áður en að tími gafst til þess að leita upplýsinga um tímaflækjur mismunandi þáttunaraðferða kom notalega á óvart.

Þegar að n hafði verið fundin var nauðsynlegt að finna d. Hún var fundin með framlengdu reikniriti Evklíðs [2] á augabragði, en þá var einkalykillinn (n, d) fundinn.

[2] Framlengda reiknirit Evklíðs finnur margföldunarandhverfu $x \pmod{n}$ sem p, sem sjá má með $px + sn = 1$. Sjá kafla 6.2

5 Afkóðunin

Eftir að d hefur verið fundin er hún sett inn í einkalykilinn (n, d). Þá er minniháttar mál að afkóða blokkirnar. Uppgefnar voru tvær blokkir C:

$$\begin{aligned} C_1 &= (00 : E68D : 5EDF : 3416 : FF31 : D5FC : 22D5 : 2E24 : 26DC \\ &\quad : 5141 : 8733 : 85D4 : 18C7 : AA7E : 1368 : 3D23 : 332F)_{16} \\ C_2 &= (01 : 7ACD : 7355 : 13A8 : 5617 : B44C : 69EA : 00DD : 8870 \\ &\quad : 54C4 : 08AF : 6E02 : CEDA : 268E : D44E : 7E3A : 91CB)_{16} \end{aligned}$$

sem afkóðaðar eru með $C^d \equiv T \pmod{n}$ eins og áður kom fram. Vegna þess hve stórar tölur það eru, þá var mjög erfitt að reikna þessi veldi í heilu lagi, en hægt er að reikna þetta með snörum hætti í tvenndarformi, og niðurstaðan var:

$$\begin{aligned} T_1 &= 357223379232748308428134812495137945540 65 \\ &\ldots\ 774225243549532524159404882405781608 \\ T_2 &= 105813664676006111832821162032459130477571 \\ &\ldots\ 177663097601224838262784579369705504 \end{aligned}$$

eða í Hexadecimal framsetningu:

$$\begin{aligned} T_1 &= (4EFA : 2073 : 6B61 : 6C20 : 6E65 : 666E : 6120 : 736F : 6E75 \\ &\quad : 204E : 6AE1 : 6C73 : 2E20 : 536B : 6172 : 7068)_{16} \\ T_2 &= (E9F0 : 696E : 6E20 : 68E9 : 7420 : 6869 : 6E6E : 2065 : 6C73 \\ &\quad : 7469 : 2E20 : 2020 : 2020 : 2020 : 2020 : 2020)_{16} \end{aligned}$$

Í ASCII framsetningu er textinn því

```
"Nú␣skal␣nefna␣sonu␣Njáls.␣Skarphéðinn
hét␣hinn␣elsti.␣␣␣␣␣␣␣␣␣␣␣"
```

(með bilunum) sem er fengið úr 25. kafla í Brennu-Njáls sögu. Klásan er:

Nú skal nefna sonu Njáls. Skarphéðinn hét hinn elsti. Hann var mikill maður vexti og styrkur, vígur vel, syndur sem selur, manna fóthvatastur, skjótráður og öruggur, gagnorður og skjótorður en þó löngum vel stilltur. Hann var jarpur á hár og sveipur í hárinu, eygður vel, fölleitur og skarpleitur, liður á nefi og lá hátt tangarðurinn, munnljótur nokkuð og þó manna hermannlegastur.

5.1 Aðrar niðurstöður

Uppgefnar stærðir:

$$\begin{aligned} n &= 22187327248325494514192235318691078870312021700746 \\ &... \ 359253895628114060254106456 1 \\ e &= 65537 \end{aligned}$$

Reiknaðar stærðir:

$$\begin{aligned} p &= 34830925363169860816412549967967 1052581 \\ q &= 63700079791122410320223306087272681 5581 \\ \phi(n) &= (p-1)(q-1) \\ &= 22187327248325494514192235318691078870213490695592 \\ &... \ 06698275899225800501431964 00 \\ de &\equiv 1 \pmod{\phi(n)} \\ &\equiv 1 \pmod{22187327248325494514192235318691} \\ &... \ 78870213490695592066982758992258005014319640 0) \end{aligned}$$

Þá sést að

$$\begin{aligned} d &= \frac{k \cdot \phi(n) + 1}{e}, k \in \mathbb{Z} \\ d &= 16886703436936015782960902966207043563883743776738 \\ &... \ 82388727006933227474730707 3 \end{aligned}$$

6 Hálfkóðar reiknirita og skilgreiningar falla

Hér fylgir útlistun á hálfkóða (*pseudocode*) fyrir þau reiknirit sem ég notaðist við, ásamt þeim áhugaverðari föllum sem ég notaði við gerð þessa verkefnis. Sleppt er reikniritum fyrir Elliptical Curve Method (ECM), vegna flækju (sjá heimildaskrá), og RSA, sem gefin er upp í kafla 2.

6.1 Eulers ϕ fallið

Totient fallið $\phi(n)$:

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

.

6.2 Fast modular exponantiation

```
procedure fastmod(a,b,n):
begin
        var result := 1;
        B := binary(b);
        k := bits in B;
        for i = 1 to k:
        begin
                result := result^2 % n;
                if (bit(B,i) = 1) then:
                        result := (result*a) mod n;
        end
        return result;
end
```

6.3 Framlengt reiknirit Evklóðs

```
procedure exteuler(a, b):
begin
        if (b = 0) then:
                return {1, 0}
        else
        begin
                temp := exteuler(b, a mod b)
                x := first(temp)
                y := last(temp)
                return {y, x-y(a div b)}
        end
end
```

7 Heimildir og ítarefni

1. The Code Book; Simon Singh; Anchor 2000; ISBN 0-38-549-5323

2. Codes, Ciphers, and other Cryptic and Clandestine Communications; Fred B. Wrixon; Black Dog & Leventhal Publishers 1998; ISBN 1-57-912-0407

3. Mathworld, Wolfram research, http://mathworld.wolfram.com

4. Wikipedia, frjálsa alfræðiorðabókin, http://en.wikipedia.org

5. GMP-ECM, http://www.komite.net/laurent/soft/ecm/ecm-6.0.1.html