Normaforma egyenletek alkalmazása a kriptográfiában

doktori (PhD) értekezés tézisei

Ködmön József

Debreceni Egyetem
Természettudományi Kar

Debrecen, 2004

Bevezetés

Az értekezés négy fejezetből áll.

Az elsőben a jelölt [12] könyve alapján rövid összefoglalást adunk a modern - bonyolultságelméleti alapokon kidolgozott - kriptográfia alapfogalmairól. Érintjük az egészségügyi informatikai vonatkozásokat is (lásd [8], [13], [14] és [11]).

A második fejezetben elsősorban a [21], [17] és [10] könyvek alapján összefoglaljuk az egyirányú függvények legfontosabb tudnivalóit.

Lényeges megjegyezni, hogy nem ismert *matematikai bizonyítás* arra, hogy ezek az egyirányú függvények valóban léteznek. Elterjedt használatuk létjogosultságát csupán a gyakorlati tapasztalatok igazolják. Helyesebb lenne az egyirányú függvény jelölt elnevezés, de mégis inkább az irodalomban szokásos egyirányú függvény szóhasználatnál maradunk.

A harmadik fejezet tartalmazza értekezésünk új eredményeit. Egy *új egyirányú hash függvény* elkészítését írjuk le, amihez a diofantikus egyenletek elméletét alkalmazzuk. Célunk eléréséhez azonban nem az egyenletek megoldásainak megtalálásához vezető algoritmusokkal foglalkozunk, hanem - éppen ellenkezőleg - az olyan egyenleteket igyekszünk alkalmazni, amelyek megoldása nem látszik kivitelezhetőnek. Egyirányú függvényünk invertálási nehézségét az algebrai számelméletből jól ismert általános normaforma egyenletek megoldásának nehézségére alapozzuk.

Új egyirányú függvényünk segítségével készítünk egy *egyirányú hash függvényt*, amelyet ajánlunk a gyakorlati felhasználás számára.

Értekezésünk fő eredménye, hogy bizonyítjuk (lásd még [3]), hogy az $\mathcal{N}_{P,s}$ függvény ütközésmentes, ha a modulus elegendően nagy. Ez pedig azt jelenti, hogy alkalmas egyirányú hash függvénynek.

A negyedik fejezetben találhatók a MAPLE rendszer-

ben készített alkalmazások forráskódjai, amelyeket az általunk - szintén MAPLE-ben - készített algoritmusok teszteléséhez használtunk.

Egyirányú függvények

Az egyirányú függvényekkel kapcsolatos fogalmak pontos definícióját adjuk meg, felhasználva a [17] és [10] könyveket.

1. Definíció. *A $\nu : \mathcal{N} \to \mathbb{R}$ függvény **elhanyagolható**, ha minden $c \geq 0$ konstanshoz létezik k_c egész úgy, hogy minden $k \geq k_c$ esetén $\nu(k) < k^{-c}$.*

2. Definíció. *Az $f : \{0,1\}^* \to \{0,1\}^*$ függvény **(erős) egyirányú** függvény ha:*
(1) Létezik PPT (polinomiális valószínűségi) algoritmus, amelynek x inputja esetén az output $f(x)$ lesz.
(2) Bármely $\mathcal{A}$ PPT algoritmushoz létezik egy $\nu_{\mathcal{A}}$ elhanyagolható függvény úgy, hogy bármely elegendően nagy k esetén:

$$\mathbf{P}\left[f(z) = y : x \overset{R}{\in} \{0,1\}^k ; y = f(x); z = \mathcal{A}(k, y)\right] \leq \nu_{\mathcal{A}}(k),$$

ahol a valószínűséget úgy vesszük, hogy az x-ek befutják lehetséges értékeiket és az $\mathcal{A}$ algoritmus is az érmefeldobásait.

Itt az $x \overset{R}{\in} \{0,1\}^k$ azt jelenti, hogy az x szót véletlenszerűen választjuk az összes lehetséges k hosszúságú bináris szavak közül.

A definícióban szereplő k konstans az úgynevezett biztonsági paraméter, amely jellemez egy adott kriptorendszert. A k paraméter általában megegyezik az x input bináris hosszával.

A szakirodalom foglalkozik másfajta egyirányú függvény definícióval is, de mi a továbbiakban ezt az erős változatot használjuk.

A gyakorlati alkalmazásokban nem egy egyirányú függvényt, hanem egy *egyirányú függvény családot* szokás használni:

3. Definíció. *Legyen I egy indexhalmaz. Minden $i \in I$ esetén D_i és R_i véges számhalmazok. Az*

$$F = \{f_i : D_i \mapsto R_i\}_{i \in I}$$

halmazt ***egyirányú függvény családnak*** *nevezzük, ha teljesíti az alábbi feltételeket:*
(1) Létezik $\mathcal{S}_1$ PPT algoritmus, ami a k inputra előállít egy legfeljebb k bináris hosszúságú $i \in I$ indexet.
(2) Létezik $\mathcal{S}_2$ PPT algoritmus, amely az $i \in I$ inputra $x \in D_i$ outputot szolgáltat.
(3) Létezik $\mathcal{A}_1$ PPT algoritmus, amely az $i \in I$ és $x \in D_i$ inputra $f_i(x)$ outputot szolgáltat, azaz $\mathcal{A}_1(i, x) = f_i(x)$.
(4) Bármely $\mathcal{A}$ PPT algoritmushoz létezik egy $\nu_{\mathcal{A}}$ elhanyagolható függvény úgy, hogy bármely elegendően nagy k esetén:

$$\mathbf{P}\left[f_i(z) = y : i \overset{R}{\in} I, \overset{R}{\in} D_i; y = f_i(x); z = \mathcal{A}(i, y)\right] \leq \nu_{\mathcal{A}}(k).$$

Ahol a valószínűséget úgy értelmezzük, hogy az i és x befutják lehetséges értékeiket és az $\mathcal{A}$ algoritmus is az érmefeldobásait.

Értekezésünkben meghatározó jelentősége van a *hash függvényeknek.*

4. Definíció. *A h függvényt* ***hash*** *függvénynek nevezzük, ha legalább az alábbi két feltételt teljesíti:*
(1) ***összenyomás****: a h függvény egy tetszőleges bináris hosszúságú x inputot egy rögzitett méretű $h(x)$ outputba képezi le.*
(2) ***könnyű kiszámítani****: adott h és x input esetén $h(x)$ kiszámítása legfeljebb polinomiális algoritmussal lehetséges.*

5. Definíció. ***Inverzkép rezisztencia****: előre megadott y outputhoz keresztülvihetetlen egy x input megadása úgy, hogy $h(x) = y$ teljesüljön.*

6. Definíció. ***Második inverzkép rezisztencia****: előre megadott x_1 inputhoz keresztülvihetetlen egy $x_2(\neq x_1)$ második input megadása úgy, hogy $h(x_1) = h(x_2)$ teljesüljön.*

7. Definíció. *Egy h hash függvény* ***ütközésmentes (collision resistant)****, ha tetszőleges x_1 és x_2 értékekre és elegendően nagy k-ra, ha $x_1 \neq x_2$, akkor*

$$\mathbf{P}\left[h(x_1) = h(x_2)\right] \leq \nu(k),$$

ahol a valószínűséget úgy értelmezzük, hogy az $x_1, x_2 \in \{0,1\}^k$ értékek egymástól függetlenül befutják lehetséges értékeiket.

8. Definíció. *(Strict avalanche criterion) Az f függvény* ***erős lavina hatással*** *rendelkezik, ha egy input bit megváltoztatása az összes output bit változását okozza $\frac{1}{2}$ valószínűséggel.*

Ez lényegében azt jelenti, hogy egy input bit változása esetén az output bitek átlagosan fele megváltozik. A bitek megváltozását a kódoláselméletből ismert *Hamming-távolság* segítségével lehet mérni. Ennek definíciója a következő:

9. Definíció. *(Hamming-távolság) Legyen az $x, y \in \mathbb{Z}$ számok bináris reprezentációja $x = (\xi_1, ...\xi_n)$ illetve $y = (\eta_1, ...\eta_n)$, ahol $\xi_i, \eta_i \in \{0,1\}$. Ekkor a két szám Hamming-távolsága:*

$$\varrho(x,y) = \sum_{i=1}^{n} \xi_i \oplus \eta_i,$$

ahol $\oplus$ jelenti a bitenként vett kizáró vagy műveletet.

A továbbiakban a közismert egyirányú függvények leírásával, típusaival és azok alkalmazásával, valamint a feltörési lehetőségekkel foglalkozunk.

Normaforma egyenlet alkalmazásán alapuló egyirányú függvény

Az egyirányú függvények kutatásának igen kiterjedt irodalma van, amelyek közül csak néhány fontosabbat emelünk most ki: [9], [5], [4], [18] és [25].

A normaforma egyenlet alkalmazásán alapuló egyirányú függvény kiszámításához három módszert hasonlítunk össze.

Ezek az $\mathcal{N}$, $\mathcal{N}_P$ és $\mathcal{N}_{P,s}$ leképezések, melyeknek sorrendisége mutatja a konstrukció fejlődését. A gyakorlatban való használatra az $\mathcal{N}_{P,s}$ függvényt ajánljuk, amely az $\mathcal{N}_P$ függvény moduláris aritmetikát használó változata.

Legyen θ egy n-ed fokú algebrai egész, és jelölje $T(X) \in \mathbb{Z}[X]$ a $\mathbb{Q}$ fölötti minimál polinomját. Legyen $K := \mathbb{Q}(\theta)$, és jelölje $\sigma_i : K \to \mathbb{C}$ a K számtest különböző beágyazásait a koplex számtestbe. Továbbá, valamely $\alpha \in K$ esetén jelölje $\alpha^{(i)} := \sigma_i(\alpha)$ $(i = 1, \ldots, n)$ az α konjugáltjait.

Legyenek $\alpha_1, \alpha_2, ..., \alpha_m \in K$ $\mathbb{Q}$-linearisan független algebrai egészek. Tekintsük az

$$L(\underline{X}) = \alpha_1 X_1 + ... + \alpha_m X_m$$

lineáris formát, ahol $m \leq n$ és legyen

$$L^{(i)}(\underline{X}) = \alpha_1^{(i)} X_1 + ... + \alpha_m^{(i)} X_m.$$

A

$$Norm_{K/\mathbb{Q}}(L(\underline{X})) = \prod_{i=1}^{n} L^{(i)}(\underline{X})$$

polinomot normaformának nevezzük. Könnyen látható, hogy $Norm_{K/\mathbb{Q}}(L(\underline{X}))$ egy n-ed fokú, egész együtthatós homogén polinom.

Tekintsük most azt a speciális esetet, amikor $\alpha_1 = 1, \alpha_2 = \theta, ..., \alpha_m = \theta^{m-1}$. Feltételezzük még, hogy $1, \theta, ..., \theta^{m-1}$ egy

hatvány egész bázisa $\mathbb{Z}_K$-nak, ahol $\mathbb{Z}_K$ jelöli a K számtest egészeinek gyűrűjét. Így minden normaforma racionális együtthatós lineáris transzformációval konvertálható az alábbi speciális alakba:

$$\mathcal{N}(\underline{X}) = Norm_{\mathbb{K}/\mathbb{Q}}(X_1 + \theta X_2 + \cdots + \theta^{m-1} X_m). \tag{1}$$

Így az általánosság megsértése nélkül használhatjuk ezt a formát.

Leképezésünk első változata az alábbi:.

1. Konstrukció. *Definiáljuk az* $\mathcal{N} : \mathbb{Z}^m \to \mathbb{Z}$ *leképezést az alábbi módon:*

$$\mathcal{N} : (x_1, \ldots, x_m) \mapsto Norm_{\mathbb{K}/\mathbb{Q}}(x_1 + \theta x_2 + \cdots + \theta^{m-1} x_m).$$

Leképezésünk második változatában legyen $P(X) \in \mathbb{Z}[X]$ rögzített n-ed fokú ($n \geq 3$) főpolinom, amelynek nincsenek többszörös gyökei. Jelölje $\alpha_1, \ldots, \alpha_n$ a P polinom gyökeit és legyen

$$L^{(i)}(\underline{X}) := \sum_{j=1}^{m} \alpha_i^{j-1} X_j, \quad \text{ahol } \; i = 1, \ldots, n \text{ és } m \leq n.$$

Legyen definiálva a P polinomhoz tartozó normaforma az alábbi módon:

$$\mathcal{N}_P(\underline{X}) := \prod_{i=1}^{n} L^{(i)}(\underline{X}).$$

Valójában $\mathcal{N}_P(\underline{X})$ a normaforma egy általánosítása, egy speciális széteső forma.

2. Konstrukció. *Definiáljuk az* $\mathcal{N}_P : \mathbb{Z}^m \to \mathbb{Z}$ *leképezést a következő módon:*

$$\mathcal{N}_P : (x_1, \ldots, x_m) \mapsto \mathcal{N}_P(x_1, \ldots, x_m). \tag{2}$$

Mivel $\mathcal{N}_P(\underline{X})$ egy egész együtthatós homogén polinom, ezért az $\mathcal{N}_P(\underline{x})$ függvényérték minden $\underline{x} \in \mathbb{Z}^m$ esetén racionális egész lesz.

Kriptográfiai szempontból jobbnak tűnik véges testet alkalmazni. Legyen valamely $s \in \mathbb{Z}$ esetén $\mathbb{Z}_s := \mathbb{Z}/s\mathbb{Z}$. Így leképezésünk harmadik változata az alábbi:

3. Konstrukció. *Legyen $s := pq$, ahol p és q megfelelően nagy prímszámok. Definiáljuk az $\mathcal{N}_{P,s} : \mathbb{Z}_s^m \to \mathbb{Z}_s$ leképezést az alábbi módon:*

$$\mathcal{N}_{P,s} : (x_1, ..., x_m) \mapsto \mathcal{N}_P(x_1, \ldots, x_m) \mod s.$$

A következőkben megmutatjuk, hogy az $\mathcal{N}(\underline{x})$, $\mathcal{N}_P(\underline{x})$ és $\mathcal{N}_{P,s}(\underline{x})$ függvényértékek kiszámítása könnyű, ha $\underline{x} \in \mathbb{Z}^m$.

Három különböző algoritmust mutatunk be a függvényértékek kiszámítására. Az $\mathcal{N}(\underline{x})$ érték kiszámítására fogunk koncentrálni, mivel az ott alkalmazott algoritmusok és a hozzájuk tartozó bonyolultsági értékek könnyen átvihetők a másik két függvényre.

A függvényérték könnyű kiszámíthatósága azt jelenti, hogy van olyan polinomiális bonyolultságú algoritmus, amellyel a függvényérték egyértelműen meghatározható. A bonyolultságot minden esetben az input értékek $\mathbb{X}$ maximuma és az n fokszám függvényében adjuk meg.

A norma definíciója alapján nem lehet egész aritmetikával számolni, ezért meghatározzuk a közelítő számítás hibájának felső korlátját is. A bonyolultságra pedig az alábbi eredményt kapjuk:

1. Tétel. [2] *Az $\mathcal{N}(\underline{x})$ függvényérték meghatározásának bonyolultsága az (1) definíció alapján $O\left(n^6 + n^4 \log^2 \mathbb{X}\right)$, ahol a O konstansa csak A értékétől függ, ahol $A = \sum_{j=1}^{m} \boxed{\alpha_j} \geq 2$ és $\boxed{\alpha_j} = \max\left\{\left|\alpha_j^{(i)}\right| : 1 \leq i \leq n\right\}$.*

A függvényértékek azonban meghatározhatók egész aritmetikával is. Ehhez az $L^{(i)}(\underline{X})$ lineáris formák mátrix reprezentációját alkalmazzuk. Bizonyítjuk a következő tételt:

2. Tétel. [3] *Legyen P egy n-ed fokú főpolinom, amelynek nincsenek többszörös gyökei. Jelölje α a P egy gyökét. Legyen $\underline{x} \in \mathbb{Z}^m$ és legyen $\mathcal{N}_P(\underline{x})$ definiálva (2) szerint. Akkor az $\mathcal{N}_P(\underline{x})$ függvényérték egész aritmetikával meghatározható.*

1. Megjegyzés. *A 2. Tétel igaz az (1)-ben definiált $\mathcal{N}(\underline{x})$ függvényérték kiszámításának esetére is. Legyen $T(x) = P(x)$ és legyen például $\theta = \alpha_1$, ekkor a tételben leírt kiszámítási módszer lényegében ugyanúgy használható.*

Az algoritmus bonyolultságára pedig a következő eredményt kapjuk:

3. Tétel. [3] *Ha az egész aritmetikát alkalmazó algoritmust használjuk az $\mathcal{N}_P(\underline{x})$ függvényérték kiszámítására, akkor annak bonyolultsága $O(n^7 + n^6 \log \mathbb{X} + n^5 \log^2 \mathbb{X})$, ahol a O konstansa csak a $P(X)$ polinom együtthatóitól függ.*

2. Megjegyzés. *Tekintettel a 1.Megjegyzésre, a 3.Tételben megadott bonyolultság igaz a (1)-ben definiált $\mathcal{N}(\underline{x})$ függvényérték kiszámításának bonyolultságára is.*

A függvényértékek meghatározhatók moduláris aritmetikával is, amelyben minden $x \in \mathbb{Z}$ számhoz, amire $-\left[\frac{M-1}{2}\right] \leq x \leq \left[\frac{M}{2}\right]$ teljesül, egy v elemű vektort tudunk rendelni a következő módon:

$$\psi(x) = x^{(M)} = (x \bmod m_1, ..., x \bmod m_v),$$

ahol $M := m_1 m_2 \cdots m_v$ és az $m_1, ..., m_v > 0$ számok páronként relatív prímek. Az $x^{(M)}$ vektort szokás az x egész szám moduláris reprezentációjának nevezni. Mivel a ψ: $x \leftrightarrow x^{(M)}$ leképezés egy gyűrű homomorfizmus, az egész aritmetikát használó algoritmusunk átalakítható úgy, hogy a moduláris reprezentáció alkalmazásával tovább csökkentjük a bonyolultságot. Az alábbi eredményeket kapjuk:

4. Tétel. [2] *Az $\mathcal{N}(\underline{x})$ függvényérték moduláris aritmetikával történő meghatározásának bonyolultsága $O(n^7 + n^6 \log \mathbb{X} + n^2 \log^{3/2} \mathbb{X})$, ahol a O konstansai csak a $T(X)$ polinom együtthatóitól függnek.*

5. Tétel. [3] *Az $\mathcal{N}_P(\underline{x})$ függvényérték moduláris aritmetikával történő meghatározásának bonyolultsága $O(n^7 + n^6 \log \mathbb{X} + n^2 \log^{3/2} \mathbb{X})$, ahol a O konstansa csak a $P(X)$ polinom együtthatóitól függ.*

Az $\mathcal{N}_{P,s}(\underline{x})$ függvényérték kiszámítására a következő eredményt kapjuk:

6. Tétel. [3] *Az $\mathcal{N}_{P,s}(\underline{x})$ függvényérték kiszámításának bonyolultsága $O(n^5 \log^2 s)$, ahol a O konstansa abszolút.*

Elkészítjük mindhárom algoritmus implementációját MAPLE rendszerben, és teszteljük a futásidőket. Ezek azt mutatják, hogy az $\mathcal{N}(\underline{x})$, $\mathcal{N}_P(\underline{x})$ és $\mathcal{N}_{P,s}(\underline{x})$ függvényértékek kiszámítása - a gyakorlatban is - valóban könnyű. A legjobbnak a mátrix reprezentációt alkalmazó algoritmus bizonyult.

Meg kell azonban azt is vizsgálnunk, hogy az $\mathcal{N}$, $\mathcal{N}_P$ és $\mathcal{N}_{P,s}$ leképezések inverzének kiszámítása mennyire nehéz feladat. Ehhez elegendő az $\mathcal{N}_P(\underline{X}) = b$, illetve $\mathcal{N}_P(\underline{X}) = b \bmod s$ normaforma egyenletek megoldásának lehetőségeit áttekinteni, ahol $0 \neq b \in \mathbb{Z}$ és a megoldásokat az $\underline{x} \in \mathbb{Z}^m$ vektorok között keressük.

Az ilyen általános normaforma egyenletek megoldására jelenleg nem ismert a gyakorlatban is használható algoritmus, sőt még az is nyitott kérdés, hogy a probléma algoritmussal megoldható-e. Mint ismeretes, Jurij Matijaszevics [16] 1970-ben bizonyította, hogy nem létezik olyan algoritmus, amely az általános diofantikus egyenletek megoldhatóságát eldöntené.

A normaforma egyenletekkel szoros kapcsolatban álló lineáris rekurzív sorozatokról szól Cerlienco, Mignotte és Piras [6] következő sejtése:

1. Sejtés. *Létezik k pozitív egész szám és $\xi^{(1)}, \ldots, \xi^{(k)}$ egészekből álló lineáris rekurzív sorozat a következő tulajdonsággal: Algoritmussal nem dönthető el, hogy vannak-e olyan $n_1, \ldots, n_k \in \mathbb{N}^k$ értékek, amelyekre $\xi_{n_1}^{(1)} + \cdots + \xi_{n_k}^{(k)} = 0$.*

Ezt felhasználva, talán belátható lenne, ha igaz ez a sejtés, akkor a normaforma egyenletek megoldhatósága algoritmikusan nem dönthető el. Pontosabban, kimondható a következő sejtés:

2. Sejtés. *Ha az 1. Sejtés igaz, akkor nem létezik olyan algoritmus, amelyik nem elfajuló normaforma egyenletekről eldönti, hogy megoldhatóak-e.*

Megvizsgáljuk a Thue-egyenlet esetét, amely tekinthető egy speciális normaforma egyenletnek. Legyen $F(X, Y) \in \mathbb{Z}[X, Y]$ irreducibilis polinom és $\deg(F) \geq 3$. Ekkor minden nem nulla $b \in \mathbb{Z}$-re az

$$F(X, Y) = b \tag{3}$$

diofantikus egyenletet Thue-egyenletnek nevezzük, ahol a megoldások $(x, y) \in \mathbb{Z}^2$ számpárok. Látható, hogy a Thue-egyenlet lényegében az $\mathcal{N}_P(\underline{X}) = b$ normaforma egyenlet két ismeretlenes specializálása ($m = 2, n \geq 3$).

A. Thue 1909-ben a [23] cikkben bizonyította, hogy (3)-nak csak véges sok megoldása van. 1968-ban A. Baker [1] effektív felső korlátot határozott meg a megoldások abszolút értékének maximumára, ami annyit jelent, hogy a (3) Thue egyenlet megoldásai algoritmussal meghatározhatóak. Huszonegy évvel később, 1989-ben Tzanakis és de Weger [24] gyakorlatban használható algoritmust adott a (3) egyenlet megoldásainak megkeresésére az LLL algoritmus (lásd [7] 2.6.3. Algoritmus) felhasználásával.

Thue-egyenletek gyakorlati megoldásának esetére N.P. Smart [22] elvégezte a bonyolultsági elemzést, amelyből kiderül, hogy ilyen egyenletek megoldása általában exponenciális bonyolultságú.

A fentiek alapján látható, hogy az $\mathcal{N}_P(\underline{x}) = b$, *illetve* $\mathcal{N}_P(\underline{x}) = b \bmod s$ *egyenlet megoldásainak meghatározása elég nehéz probléma. Így az* $\mathcal{N}$, $\mathcal{N}_P$ *és* $\mathcal{N}_{P,s}$ *függvények lehetnek egyirányú függvény jelöltek, különösen akkor, ha a változók száma legalább 3.*

Értekezésünkben az $\mathcal{N}_P$ és $\mathcal{N}_{P,s}$ függvények két fontos tulajdonságát - az ütközésmentességet és a lavina hatást - vizsgáljuk meg.

Az $\mathcal{N}_{P,s}$ függvény ütközéseinek bekövetkezési valószínűségét a

$$P_{coll} = P[\mathcal{N}_{P,s}(\underline{x}) = \mathcal{N}_{P,s}(\underline{y}) : \underline{x}, \underline{y} \in {\mathbb{Z}_s}^m]$$

kifejezés értéke mutatja meg.

Értekezésünk fő eredménye a következő két tétel:

7. Tétel. [3] *Legyen* $P(X) \in \mathbb{Z}[X]$ *egy legalább harmadfokú főpolinom, amelynek nincsenek többszörös gyökei. Legyenek* p *és* q *olyan prímszámok, hogy* $q > p > q/2$ *és* $s := pq$. *Tegyük fel, hogy* $lnko(m, \varphi(s)) = 1$. *Jelölje* $N(P, b, s)$ *az* $N_P(x_1, \ldots, x_m) \equiv b \pmod{s}$ *kongruencia megoldásainak számát.*

- *Ha* $lnko(b, s) = 1$ *teljesül, akkor*

$$|N(P, b, s) - s^{m-1}| < c_1(P) s^{m-1-\frac{1}{4}}$$

- *egyébként pedig*

$$N(P, b, s) < c_2(P) s^{m-1}.$$

A tételben $lnko(n, m)$ jelöli az n és m egészek legnagyobb közös osztóját, $\varphi(x)$ pedig, szokás szerint, az Euler féle φ-függvényt.

8. Tétel. [3] *Legyen* $P(X) \in \mathbb{Z}[X]$ *egy legalább harmadfokú főpolinom, amelynek nincsenek többszörös gyökei. Legyenek* p *és* q *prímszámok olyanok, hogy* $q > p > q/2$ *és* $s := pq$. *Tegyük*

fel, hogy $lnko(m, \varphi(s)) = 1$. *Akkor az* $\mathcal{N}_{P,s}$ *függvény* P_{coll} *ütközési valószínűsége teljesíti a*

$$P_{coll} < \frac{C}{s},$$

egyenlőtlenséget, ahol a C *konstans csak a* P *polinomtól függ.*

A bizonyításhoz alkalmazzuk S. Lang és A. Weil [15] tételét, amely szerint ha a p prím modulus elegendően nagy, akkor az $f \equiv 0 \pmod{p}$ egyenlet $N(f, p)$ megoldásainak száma teljesíti az

$$\left|N(f,p) - p^{m-1}\right| < C(f)p^{m-1-\frac{1}{2}}$$

egyenlőtlenséget, ahol a $C(f)$ konstans értéke kizárólag az $f(X_1, \ldots, X_m)$ abszolút irreducibilis polinomtól függ.

Az abszolút irreducibilitás bizonyításához két lemmára lesz szükségünk. Az első a klasszikus Schöneman - Eisenstein tétel megfelelője polinom gyűrűkre és egy következménye Capelli tételének (lásd [20] 662. oldal).

1. Lemma. *Legyen* $P(X) \in \mathbb{C}[X]$ *egy polinom, amelynek van legalább egy egyszeres gyöke. Akkor a* $Z^n - P(X) \in \mathbb{C}[X, Z]$ *polinom abszolút irreducibilis minden* $n \geq 1$ *esetén.*

2. Lemma. *Legyen* $f(\underline{X}) := f(X_1, \ldots, X_m) := \prod_{i=1}^{n}(\alpha_{i1}X_1 + \cdots + \alpha_{im}X_m) \in \mathbb{C}[X_1, \ldots, X_m]$ *egy forma, amelyre teljesül, hogy* $\alpha_{ij} \in \mathbb{C}$ *minden* $1 \leq i \leq n$, $1 \leq j \leq m$ *esetén, és* $\prod_{i=1}^{n}\prod_{j=1}^{m}\alpha_{ij} \neq 0$. *Továbbá tegyük fel, hogy létezik* $1 \leq j_1 < j_2 \leq n$ *úgy, hogy*

$$\frac{\alpha_{ij_1}}{\alpha_{ij_2}} \notin \left\{\frac{\alpha_{kj_1}}{\alpha_{kj_2}} : 1 \leq k \leq n, k \neq i\right\}.$$

Akkor az $f(\underline{X}) - a$ *polinom irreducibilis* $\mathbb{C}[\underline{X}]$*-ben minden* $0 \neq a \in \mathbb{C}$ *esetén.*

3. Megjegyzés. *Ha az s modulus elég nagy, akkor az $\mathcal{N}_{P,s}$ függvény a 7. Definició szerint ütközésmentes.*

Az $\mathcal{N}_{P,s}$ és $\mathcal{N}_P$ függvények lavina hatásának teszteléséhez egy MAPLE programot készítünk, amely az alábbi relatív Hamming-távolságokat határozza meg:

$$RH_{P,s} := \frac{\rho(\mathcal{N}_{P,s}(\underline{x}), \mathcal{N}_{P,s}(\underline{x}'))}{l(s)}$$

és

$$RH_P := \frac{\rho(\mathcal{N}_P(\underline{x}), \mathcal{N}_P(\underline{x}'))}{l(\mathcal{N}_P(\underline{x}'))},$$

ahol az $\underline{x}$ egy input az $\underline{x}'$ pedig a hozzá tartozó egy bitben megváltoztatott input, valamint $l(.)$ jelöli a bináris hossz függvényt és $\rho(.,.)$ a 9. Definicióban megadott Hamming-távolságot.

A tesztet több alkalommal is ezerszer futtattuk véletlenszerűen generált 2^{1024} nagyságrendű input adatokkal. Az $\underline{x}'$ input egyetlen bitjének megváltoztatása szintén véletlenszerű volt.

Az $RH_{P,s}$ értékek átlaga 0.50 volt, 0.004 szórással. Az $\mathcal{N}_P$ függvény tesztje egy kicsit rosszabb eredményt hozott, az RH_P értékek átlaga 0.48 volt, 0.003 szórással. Ez a függvény csak közelítőleg teljesíti az 8. Definícióban megfogalmazott erős lavinahatás kritériumot.

3. Sejtés. *Az $\mathcal{N}_{P,s}$ függvény teljesíti az 8. Definícióban megfogalmazott erős lavinahatás kritériumot.*

Azt, hogy az $\mathcal{N}_{P,s}$ függvény invertálása általában nehéz, az alábbi feltétel fejezi ki.

10. Definíció. *Erős Moduláris Normaforma Feltétel (SMNA): Minden Q polinom és $\mathcal{A}$ PPT algoritmus esetén, ha az s elegendően nagy*

$$P[\mathcal{A}(s,b) = (x_1, ..., x_m) \text{ úgy, hogy } b = \mathcal{N}_{P,s}(x_1, ..., x_m)] < \frac{1}{Q(k)},$$

ahol $x_i \in \mathbb{Z}$ és $1 \leq \max\{|x_i|\} \leq s-1$, valamint a valószínűséget úgy kell venni, hogy az x_i inputok befutják lehetséges értékeiket, és az $\mathcal{A}$ algoritmus is az érmefeldobásait.

4. Megjegyzés. *Az előbbihez hasonlóan az $\mathcal{N}_P(\underline{X})$ függvényhez lehet definiálni az Erős Normaforma Feltételt (SNA), ami az $\mathcal{N}_P$ függvény invertálásának keresztülvihetetlenségét fejezi ki.*

Bizonyítható a következő tétel.

9. Tétel. *Az SMNA, illetve SNA teljesülése esetén az $\mathcal{N}_{P,s}$, illetve az $\mathcal{N}_P$ függvény egyirányú hash függvény.*

A kriptográfiai gyakorlat általában nem egy egyirányú függvényt alkalmaz, hanem inkább *egyirányú függvény családot.* Egy ilyen családot fogunk készíteni az $\mathcal{N}_{P,s}$ függvény felhasználásával. Tekintsük az alábbi konstrukciót:

4. Konstrukció. *Legyen $P(X) \in \mathbb{Z}[X]$ egy olyan főpolinom, amelynek nincsenek többszörös gyökei. Legyenek p és q elegendően nagy prímszámok úgy, hogy $q > p > q/2$ és $s := pq$. Tegyük fel továbbá, hogy $lnko(m, \varphi(s)) = 1$. Definiáljuk a $\mathcal{P} = \{P(X) : deg(P(X)) = n \geq 4,$ rögzített $\}$ polinom halmazt és az*

$$\mathbf{MNFF} = \left\{ \mathcal{N}_{P,s} : \mathbb{Z}_s^m \mapsto \mathbb{Z}_s;\ \mathcal{N}_{P,s}(\underline{X}) = (\prod_{i=1}^{n} L_i(\underline{X}), \mathrm{mod}\, s) \right\}_{P \in \mathcal{P}}$$

függvény családot, ahol $3 \leq m \leq n$, és az $L^{(i)}(\underline{X})$ definíciója ugyanaz, mint korábban.

Az alábbi tételt nyerjük:

10. Tétel. *Ha minden $P(X) \in \mathcal{P}$ polinomhoz tartozó $\mathcal{N}_{P,s}$ függvény teljesíti az SMNA-t, akkor az MNFF függvény halmaz egy egyirányú hash függvény család.*

5. Megjegyzés. *A fentihez hasonló módon lehet az $\mathcal{N}_P$ függvény alkalmazásával megkonstruálni az NFF függvénycsaládot, amelyről bizonyítani lehet, hogy SNA esetén az NFF egy egyirányú függvénycsalád.*

Megvizsgáljuk, hogy miként kell megválasztani a $P(X)$ polinomot, hogy a lehető legegyszerűbben lehessen számolni a függvényértéket.

Legyen $P(X) := X^n - 1$ úgy, hogy $n \geq 3$. Jelölje ζ ennek egy gyökét és legyen

$$L(\underline{X}) := X_1 + \zeta X_2 + \cdots + \zeta^{n-1} X_n.$$

A

$$\zeta^j L(\underline{X}) = X_{n-j+1} + \zeta X_{n-j+2} + \cdots + \zeta^{j-1} X_n + \zeta^j X_1 + \cdots + \zeta^{n-1} X_{n-j}$$

egyenletből, amely minden $1 \leq j \leq n$ esetén fennáll, látható, hogy az $\mathcal{N}_{P,s}(\underline{X})$ egy

$$\begin{pmatrix} X_1 & X_2 & \dots & X_n \\ X_n & X_1 & \dots & X_{n-1} \\ \dots & \dots & \dots & \dots \\ X_2 & X_3 & \dots & X_1 \end{pmatrix}$$

ciklikus mátrix determinánsa, amely nagyon egyszerű formájú. Ha $L(\underline{X}) = X_1 + \zeta X_2 + \cdots + \zeta^{m-1} X_m$, ahol $m < n$, akkor $X_{m+1}, \ldots, X_n$ helyett minden sorban $n - m$ darab 0 áll, de a determináns értéke nem lesz 0, csak a formája lesz még egyszerűbb.

Végül egy rekurzív hash függvényt konstruálunk az $\mathcal{N}_{P,s}(x_1, ..., x_m)$ függvény alkalmazásával.

Legyen adott az M üzenet, amelyet egy tetszőleges hosszúságú bináris szónak tekintünk. Ezt szétvágjuk $x_1, ..., x_k$ részszavakra úgy, hogy minden x_i $(1 \leq i \leq k)$ egy egész számot reprezentáljon az $[1, s-1]$ intervallumban. Tegyük fel, hogy $k \geq m$, egyébként pedig egészítsük ki a sorozatot nullákkal. Először definiáljuk a

$$h(x_1, \ldots, x_m) := N_{P,s}(x_1, \ldots, x_m)$$

kezdőértéket, majd pedig legyen rekurzív módon

$$h(x_1,\dots,x_{m+t(m-1)}) :=$$
$$N_{P,s}(h(x_1,\dots,x_{m+(t-1)(m-1)}), x_{m+(t-1)(m-1)+1},\dots,x_{m+t(m-1)}).$$

Ha k valamely alkalmas $t \in \mathbb{Z}$ számra nem $m+t(m-1)$ alakú, akkor egészítsük ki az M üzenetet 0-val, mindaddig, míg k megfelelő alakú lesz.

Így kapunk egy gyakorlatban is használható h egyirányú hash függvényt. Ez a függvény minden iterációs lépésben megtartja ütközésmentes tulajdonságát és sejthetően rendelkezik lavina hatással is (lásd 8. Tétel és 3. Sejtés). Ezek a tulajdonságok garantálják biztonságos használatát.

Az értekezés tartalmaz még két konkrét számítási példát, amely bemutatja, hogyan használható a h függvény egyirányú hash függvényként a kriptográfiai gyakorlatban.

Summary

In this thesis we present a new one-way function with collision resistance and avalanche effect. The security of this function based on the difficulty of solving a general norm form equation.

Let $P(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree $n \geq 3$ having no multiple roots. Denote by $\alpha_1,\dots,\alpha_n$ the roots of P and put

$$L^{(i)}(\underline{X}) := \sum_{j=1}^{m} \alpha_i^{j-1} X_j \quad \text{for} \quad i = 1,\dots,n \text{ and } m \leq n.$$

Define the norm form corresponding to the polynomial P by

$$\mathcal{N}_P(\underline{X}) := \prod_{i=1}^{n} L^{(i)}(\underline{X}).$$

In fact, $\mathcal{N}_P(\underline{X})$ is a generalization of the concept of norm form and is a special decomposable form.

Define the mapping $\mathcal{N}_P : \mathbb{Z}^m \to \mathbb{Z}$ in the following way:

$$\mathcal{N}_P : (x_1, \ldots, x_m) \mapsto \mathcal{N}_P(x_1, \ldots, x_m).$$

Since $\mathcal{N}_P(\underline{X})$ is a homogeneous polynomial of degree n, with integer coefficients the function value $\mathcal{N}_P(\underline{x})$ will be a rational integer for every $\underline{x} \in \mathbb{Z}^m$.

Let p and q be primes such that $q > p > q/2$ and $s := pq$. Define the mapping $\mathcal{N}_{P,s} : \mathbb{Z}_s^m \to \mathbb{Z}_s$ in the following way:

$$\mathcal{N}_{P,s} : (x_1, ..., x_m) \mapsto \mathcal{N}_P(x_1, \ldots, x_m) \pmod{s}.$$

We present three algorithms for the calculation of the value $\mathcal{N}_P(\underline{x})$ and $\mathcal{N}_{P,s}(\underline{x})$. Their complexity are discussed and the running time of their imlementations in MAPLE are compared.

And so we proved that the value $\mathcal{N}_P(\underline{x})$ and $\mathcal{N}_{P,s}(\underline{x})$ can be computed efficiently.

Denote by $P_{coll} = P[\mathcal{N}_{P,s}(\underline{x}) = \mathcal{N}_{P,s}(\underline{y}) : \underline{x}, \underline{y} \in {\mathbb{Z}_s}^m]$ the probability of the collision for the function $\mathcal{N}_{P,s}$.

Our main result is the following two theorems.

Theorem 7 *Let $P(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree at least 3 having no multiple roots. Let p and q be primes such that $q > p > q/2$ and $s := pq$. Suppose that $gcd(m, \varphi(s)) = 1$. Let $N(P, b, s)$ denote the number of solutions of the congruence $N_P(x_1, \ldots, x_m) \equiv b \pmod{s}$.*

- *If $gcd(b, s) = 1$ we have*

$$|N(P, b, s) - s^{m-1}| < c_1(P) s^{m-1-\frac{1}{4}}$$

- *othervise*

$$N(P, b, s) < c_2(P) s^{m-1}.$$

Theorem 8 *Let $P(X) \in \mathbb{Z}[X]$ be a monic polynomial of degree at least 3 having no multiple roots. Let p and q be primes such that $q > p > q/2$ and $s := pq$. Suppose that $gcd(m, \varphi(s)) = 1$. Then the probability of collision P_{coll} for the function $\mathcal{N}_{P,s}$ satisfies the inequality*

$$P_{coll} < \frac{C}{s},$$

where the constant C depends only on the polynomial P.

We tested a MAPLE implementation of the function $\mathcal{N}_{P,s}(\underline{x})$ and we can showed that this function has avalanche effect.

We construct an application, which uses this function as a one-way hash function. This application can check passphrase of user of a computer. An other application makes the hash value of a message.

So we propose to apply the function $\mathcal{N}_{P,s}$ in the practice of the cryptography as a one-way hash function.

Further we found also some fact, which show that $\mathcal{N}_{P,s}$ is probably a family of one-way functions.

Hivatkozások

[1] A. Baker, *Contributions to theory of diophantine equations I. and II.*, Phil. Trans. Roy. Soc. London Ser A., **263** (1968) 173-208.

[2] A. Bérczes, J. Ködmön, *Methods for the calculation of values of a norm form*, Publ. Math. Debrecen, **63** (2003), 751-768.

[3] A. Bérczes, J. Ködmön, A. Pethő, *A one-way function based on norm form equations*, Periodica Math. Hungar. (megjelenés alatt).

[4] J. Black, S. Halevi, H. Krawczyk, T. Krovetz, P. Rogaway, *UMAC: Fast and Secure Message Authetication*, In Advances in Cryptology - CRYPTO '99 (1999), vol. 1666 of Lecture Notes In Computer Science, Springer-Verlag, pp. 216-233.

[5] J. Buchmann, S. Paulus, *A one-way function based on ideal arithmetic in number fields*, Lect. Notes Comput. Sci. **1294** (1997), 385-394.

[6] L. Cerlienco, M. Mignotte, F. Piras, *Suites récurrentes linéaires. Propriétés algébriques et arithmétiques. (Linear recurrent sequences. Algebraic and arithmetic properties)*, L'Enseign. Math. **33** (1987), 67-408.

[7] H. Cohen, *A course in computational algebraic number theory*, Springer Verlag, 1993.

[8] *Egészségügyi informatika*, Szerk: Kékes E., Surján Gy., Balkányi L., Kozmann Gy., Medicina Könyvkiadó, 2000.

[9] O. Goldreich, L. Levin, N. Nisan, *On constructing 1-1 one-way functions*, Electronic colloquium on computational complexity, TR-95-029, 6/25/95, 1995.

[10] S.Goldwasser, M.Bellare, *Lecture Notes on Cryptography*, MIT Press, Cambridge, Massachusetts 2001.

[11] *Handbook of medical informatics*, Eds: J. H. van Bemmel, M. A. Musen, Springer Verlag, 2002.

[12] Ködmön J., *Kriptográfia, Az informatikai biztonság alapjai, A PGP kriptorendszer használata*, ComputerBooks Könyvkiadó, 1999.

[13] Ködmön J., *Adatvédelem, adatbiztonság az egészségügyben*, Lege Artis Medicinae **4**, (1997), 278-281.

[14] Ködmön J., *Az egészségügyi adatvédelem szabályozása*, Orvosi Hetilap **140** (1999), 1113-1115.

[15] S. Lang, A. Weil, *Number of points of varieties in finite fields*, Am. J. Math. **76** (1954), 819-827.

[16] Y.V.Matiyasevich, *Hilbert's Tenth Problem*, MIT Press,1993.

[17] A.J.Menezes, P.C.van Oorschot, S.Vanstone, *Hadbook of Applied Cryptography*, CRC Press, 1997.

[18] J. Patarin, *Secret public key schemes*, In:Public-Key Criptography and Computational Number Theory, (ed. by K. Alster, J. Urbanowicz and H. C. Williams), 221-237. Walter de Gruyter, 2001.

[19] A.Pethő, *Algebraische Algorithmen*, Vieweg Verlag,1999.

[20] L. Rédei, *Algebra I*, Akademische Verlagsgeselschaft Geest & Portig K.-G., Leipzig, 1959.

[21] B. Schneier, *Applied Cryptography*, John Wiley & Sons, 1996.

[22] N.P.Smart, *How Difficult is it to Solve a Thue Equation?*, In ANTS-2, LNCS **1122**, 363-373, 1996.

[23] A. Thue, *Über Annäherungswerte algebraischer Zahlen*, J. Reine Angew. Math. **135** (1909), 284-305.

[24] N. Tzanakis, B.M.M. de Weger, *On the practical solution of the Thue equation*, J. Number Theory, **31** (1989), 99-132.

[25] D.L. Whiting, M.J. Sabin, *Montgomery Prime Hashing for Message Authentication*, CT-RSA 2003 (ED: M. Joye), LNCS **2612**, pp. 50-67, Springer-Verlag, 2003.